

DE GRUYTER

*Javier Echeverria (Ed.) et al.*

# THE SPACE OF MATHEMATICS

GRUNDLAGEN DER KOMMUNIKATION UND  
KOGNITION / FOUNDATIONS OF COMMUNICATION  
AND COGNITION

Grundlagen der Kommunikation und Kognition  
Foundations of Communication and Cognition  
Bibliotheksausgabe/Library Edition

Herausgeber/Editors  
Roland Posner, Georg Meggle



# The Space of Mathematics

Philosophical, Epistemological,  
and Historical Explorations

edited by  
Javier Echeverria  
Andoni Ibarra  
Thomas Mormann



Walter de Gruyter · Berlin · New York  
1992



∞ Gedruckt auf säurefreiem Papier, das die US-ANSI-Norm über Haltbarkeit erfüllt.

Printed on acid-free paper which falls within the guidelines of the ANSI to ensure permanence and durability

*Die Deutsche Bibliothek — CIP-Einheitsaufnahme*

The **space of mathematics** : philosophical, epistemological, and historical explorations / ed. by Javier Echeverria ... — Berlin ; New York : de Gruyter, 1992

(Foundations of communication and cognition : library edition)

ISBN 3-11-013249-4

NE: Echeverria, Javier [Hrsg.]

© Copyright 1992 by Walter de Gruyter & Co., D-1000 Berlin 30

Dieses Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Printed in Germany

Druck: Arthur Collignon GmbH, Berlin

Buchbinderische Verarbeitung: Lüderitz & Bauer, Berlin

## Table of Contents

Acknowledgments.....	VII
Introductory afterthoughts.....	IX
<i>1. Structural Dimensions.....</i>	<i>1</i>
Mac Lane, S.: <i>The Protean Character of Mathematics</i> .....	3
Lawvere, F. W.: <i>Categories of Space and of Quantity</i> .....	14
Ibarra, A. / Mormann, T.: <i>Structural Analogies Between Mathematical and Empirical Theories</i> .....	31
Rantala, V.: <i>Reduction and Explanation: Science vs. Mathematics</i> .....	47
Niiniluoto, I.: <i>Reality, Truth, and Confirmation in Mathematics – Reflections on the Quasi-empiricist Programme</i> .....	60
Breger, H.: <i>Tacit Knowledge in Mathematical Theory</i> .....	79
Grattan-Guinness, I.: <i>Structure-similarity as a Cornerstone of the Philosophy of Mathematics</i> .....	91
<i>2. Dimensions of Applicability .....</i>	<i>113</i>
Resnik, M. D.: <i>Applying Mathematics and the Indispensability Argument</i> .....	115
Torretti, R.: <i>Mathematical Structures and Physical Necessity</i> .....	132
Scheibe, E.: <i>The Role of Mathematics in Physical Science</i> .....	141
Schmidt, H.-J.: <i>The Status of Set-theoretic Axioms in Empirical Theories</i> .....	156
Da Costa, N. C. A. / Doria, F. A.: <i>Suppes Predicates for Classical Physics</i> .....	168
Howson, C.: <i>Mathematics in Philosophy</i> .....	192
<i>3. Historical Dimensions.....</i>	<i>203</i>
Dauben, J.: <i>Are There Revolutions in Mathematics?</i> .....	205
Echeverría, J.: <i>Observations, Problems, and Conjectures in Number Theory – The History of the Prime Number Theorem</i> .....	230
Knobloch, E.: <i>Historical Aspects of the Foundations of Error Theory</i> .....	253
Jahnke, H. N.: <i>A Structuralist View of Lagrange's Algebraic Analysis and the German Combinatorial School</i> .....	280

Otte, M.: <i>Constructivism and Objects of Mathematical Theory</i> .....	296
Feferman, S.: <i>Turing's "Oracle": From Absolute to Relative Computability – And Back</i> .....	314
Mahoney, M. S.: <i>Computers and Mathematics: The Search for a Discipline of Computer Science</i> .....	349
4. <i>Global Dimensions of Knowledge: Information, Implementation, and Intertheoretic Relations</i> .....	365
Mosterín, J.: <i>Theories and the Flow of Information</i> .....	367
Sneed, J. D.: <i>Structuralism and Scientific Discovery</i> .....	379
Moulines, C. U.: <i>Towards a Typology of Intertheoretical Relations</i> .....	403
Index of Names.....	413

## Acknowledgments

In September 1990 philosophers, historians of science and mathematicians gathered at the University of the Basque Country (UPV/EHU) in Donostia/San Sebastian (Basque Country, Spain) to participate in a multidisciplinary symposium on "Structures in Mathematical Theories". The conference was organized by the Department of Logic and Philosophy of Science of UPV/EHU with the collaboration of the Spanish Society for the History of Sciences and Techniques (SEHCyT). It took place in the School of Education, and was attended by more than two hundred people. The articles in this volume are revised versions of papers presented at the conference.

Since the symposium constituted the genesis of this book, it is appropriate to acknowledge those groups and individuals whose help made it possible.

Professor Dr. C.U. Moulines deserves special credit for his intellectual support and expert advice. The project of the symposium was supported by a number of different agencies and foundations. Primary financial help came from the Dirección General de Investigación Científica y Técnica which supported the seminal research project (PB-86-0222). Further generous subsidies were provided, among others, by the Basque Government, the Diputación Foral de Gipuzkoa and the UPV/EHU.

No large conference can function without a group of key people. In this case, María Jesús Arribas, Amparo Díez and Xabier Eizaguirre worked with us for more than a year on the planning. To them, we owe a special word of thanks. Further, the symposium owes a tremendous amount to the help of the excellent staff of the Department of Logic and Philosophy of Science of UPV/EHU.

The symposium was structured around a number of topics, and some of that structure has survived the transition to book form. Thirty-five invited papers and seventy-five contributed papers were read.<sup>1</sup> We acknowledge with gratitude the efforts and kind co-operation of all the participants. We wish to

---

<sup>1</sup> Contributed papers have already been published in A. Díez, J. Echeverría, A. Ibarra (Eds.), *Structures in Mathematical Theories. Reports of the San Sebastian International Symposium*, Bilbao (Spain), University of the Basque Country press, 1990.

thank especially a number of scholars who served as commentators. It was a difficult task to select the papers to be published together as the conference volume: we have aimed to maintain a thematic coherence, somewhat stricter than was provided by the very broad spectrum of the conference itself. Hence, sometimes, we were forced to exclude outstanding papers because they did not fit nicely into the frame.

The preparation of the final manuscript that resulted from all these efforts was no less onerous a task than the organization of the symposium itself. Our principal debt here again is to Amparo Díez who was chiefly responsible for entering a vast and unhomogenous bunch of manuscripts into the computer. She kept the computer running, and her cheerful, thoughtful, and very careful assistance made the production of the book run much more smoothly than we had any right to hope for. Elena Aranda's assistance with editorial matters has also been invaluable, and we are very grateful to her as well.

Since none of us is a native speaker of English, we essentially depended on the linguistic assistance of our translators and interpreters David Simon and Diana Lindsay. It has been a pleasure to work with them. Finally we are greatly indebted to our publisher de Gruyter who treated the project with great patience and provided us throughout with useful technical advice. Thanks!

The Editors

## *Introductory afterthoughts*

1. Writing a preface for this volume resembles the activity of Minerva's owl that starts its flight at dusk: after the days of lively lectures and intensive discussions from which this volume arose have passed, the editors lay claim to the doubtful privilege of adding some final introductory afterthoughts in order to make out a (necessarily) grey and dim *philosophical gestalt* of what has taken place during the day. However, the temptation to try to get at least some idea of what is going on in the field of philosophy of mathematics in general, today, is almost irresistible: presently, philosophy of mathematics is replete with optimistic metaphors and programmatic announcements that promise a fresh start to the whole discipline. Something important, it seems, is going on in the field, and there are new developments that drastically change our perspective. Hence it might be justified to tentatively sketch a conceptual space where the contributions of this volume can be located in such a way that a pattern or *gestalt* becomes visible. Of course, complex phenomena can be perceived under different *gestalts* and it might well be the case that the reader perceives other *gestalts* which allow him to assess in a better way how the pieces of this collection fit (or do not fit) together with some of the allegedly general trends of the "New" philosophy of mathematics. In any case, in these introductory remarks we do not want to comment individually on each contribution of this volume. Rather we point to one or two salient features of each of them in order to give the reader some general idea of the whole enterprise.

2. As is evidenced by a wealth of recent publications the philosophy of mathematics presently is undergoing a rather dramatic transformation and reorientation<sup>1</sup>. What has happened to justify such optimism? Have all the

---

<sup>1</sup> To mention just a few: a recent volume of *Synthese*, edited by R. Hersh, is exclusively dedicated to "New Directions in the Philosophy of Mathematics" (not to be confused with Tymoczko's anthology of the same title from 1985). There we find contributions like Goodman's "Modernizing the Philosophy of Mathematics" or Maddy's "Philosophy of Mathematics: Prospects for the 1990's". Of the same vintage but somewhat older are Lakatos' "Renaissance of Empiricism in Recent Philosophy of Mathematics" (1967), and Hersh's "Proposals for Reviving the Philosophy of Mathematics" (1979).

problems of the traditional philosophy of mathematics been solved, and what might be the new agenda? These are the questions we want to touch upon in the following few pages.

Since its inception mathematics has been a constant source of thought and reflection for very different kinds of people: apart from mathematicians who intended to explicate their science to other scientists or to the layman, from the beginning we find philosophers who were interested in mathematics not for *internal* mathematical reasons, but for *external*, more far reaching and more ambitious reasons. For instance, Plato claimed the study of mathematics to be an essential preliminary to philosophy. For him mathematical knowledge was the paradigmatic model of real knowledge (episteme) contrasted with mere opinion (doxa) that all other cognitive enterprises (except philosophy and mathematics) could obtain. A revival of the Platonist programme took place in the 17th century when Descartes, Leibniz, and Spinoza took mathematics as *the* model of scientific knowledge - see Howson's contribution on the role of mathematics in philosophy. Later, historians, psychologists, sociologists embarked on often bold and far reaching conceptual explorations into the formidable and labyrinthine space of mathematics in order to underpin various theses on the development, function, and structure of human cognition in general.

It cannot be said that all these enterprises have been successful, often they have been endeavoured with insufficient equipment and preparation resulting in rather inadequate and distorting maps of the space of mathematics. Be that as it may, from antiquity to our days mathematics has served as a kind of guinea pig or touchstone for quite a variety of general philosophical accounts of what the world and our knowledge of it is like.

3. In the early and middle 20th century this model character of mathematics has been perceived of by professional philosophers and historians of science in a way that has emphasized the idiosyncratic character of mathematical cognition thereby isolating it from the related disciplines such as philosophy and history of the other sciences. Many philosophers and historians have even doubted that mathematics has a genuine history comparable to, say, the history of literature. They simply considered it as "an intellectual field in which historical development is swallowed up by the latest state of the art, at the same time preserving what remains worthwhile" (Otte). Although such a view of disciplinary history in the case of science has long been discredited it has had its adherents in the case of history of mathematics till very recently. For example, the historian of mathematics M. Crowe formulated ten "laws" concerning the history of mathematics; one of them (No. 10) stated that

revolutions never occur in mathematics. Dauben in his contribution explicitly argues against Crowe's antirevolutionary law<sup>2</sup>.

As is shown in the contributions of Knobloch (Foundations of error theory), Jahnke (Structure of mathematical theories according to the Combinatorial school of 19th century) even such seemingly "platonic" topics such as foundations and structure of mathematical theories are deeply impregnated by conceptual, pragmatic and historical considerations. As is shown by Feferman for recursion theory, and by Mahoney for Computer Science, this is true as well for modern theories. Once again the contributions of this volume provide ample evidence for Lakatos' Kantian slogan "philosophy of mathematics without history is empty, history of mathematics without philosophy is blind", the historical contributions implicitly or explicitly deal with philosophical issues, and the philosophical contributions rely on historical evidence to support the theses they present. This concern with historical and sociological aspects of mathematical cognition clearly separates the "New" philosophy of mathematics from the traditional Neo-Fregean approach that has dominated the discipline in the first half of this century.

4. The main concern of the Neo-Fregean approach has been the problem of secure and immutable foundations of mathematics. This has separated mathematics from other cognitive enterprises, and especially those linked to empirical knowledge which obviously were not built on bedrock, as we all know in the age of fallibilism. Even if philosophers of mathematics (as philosophers in general) in no way agreed upon how such a foundation could be achieved, the competing schools of Logicism, Formalism, and Intuitionism all adhered to the foundational task. This has lead to distorting effects for the whole discipline. Often, philosophy of mathematics has been content with an elementarist account of mathematics restricted to some known examples of elementary theories, say, arithmetic of natural numbers or basic Euclidean geometry. Then the question why  $2 + 2 = 4$  may seriously trouble the traditional philosopher of mathematics. Somehow, this looks funny. Of course, philosophers are otherworldly people and see problems where ordinary people do not, and indeed there may be deeply hidden philosophical problems in the elementary arithmetic truth  $2 + 2 = 4$  that escape the mathematician, the scientist, as well as the layman. However, philosophy of mathematics cannot restrict itself to these kinds of questions. Today, there is virtual unanimity that

---

<sup>2</sup> It might be interesting to note that also Crowe, with some reservations, no longer denies occurrence of revolutions in mathematics (cf. Aspray/Kitcher 1988, 260ff).



all the above mentioned traditional approaches have turned out to be dead ends. Their highlights and their shortcomings are well known and well documented<sup>3</sup>. In this volume the reader will find virtually nothing about these sorts of things. All contributions of this book belong, for better or for worse, to what is still called in Aspray/Kitcher (1988) the "maverick tradition" and nowadays seems to form the new mainstream of philosophy of mathematics.

We do not want to repeat yet again any of the well-known details of the traditional philosophy of mathematics. Instead of this we would like to sketch in a few sentences the general frame that has determined the various currents of the Fregean approach. The preoccupation with the problem of foundations on the one hand and the presupposition of unquestioned empiricist premises on the other hand have led the traditional approaches of philosophy of mathematics into a trap well known as "Benacerraf's dilemma":

If mathematics is the study of projective, ideal entities without position in space or time how can mankind, being confined so obviously to a tiny portion of space and time, manage to have any mathematical knowledge?

For the non-philosopher, mathematician or layman, Benacerraf's dilemma probably looks like a philosophical artifact which has an air of exaggeration and absurdity: Without doubt we have a lot of mathematical knowledge, and looking more closely at how this knowledge is manufactured one realizes it is not as special as some philosophers (and philosophically inclined mathematicians) might have us believe. Assertions claiming the "ideal" and the "absolute" character of mathematical knowledge should not be taken too seriously. After all, mathematical knowledge is human knowledge. Thus, the general strategy to overcome Benacerraf's dilemma is rather clear: one has to show

- (1) The phenomena mathematical knowledge deals with are not so ideal, i.e. they are not so remote and inaccessible, as it at first sight might appear;
- (2) One has to dismiss a too simplistic (empiricist) theory of knowledge that restricts the realm of knowable phenomena to those we can have a causal relation with<sup>4</sup>.

Today, a growing number of philosophers of mathematics consider this to be a promising strategy to escape Benacerraf's trap. This has the sobering effect of knocking mathematical knowledge off the pedestal on which Plato put it and

---

<sup>3</sup> For a short but authoritative history of Neo-Fregean philosophy of mathematics see the "opinionated Introduction" of Aspray/Kitcher (1988).

<sup>4</sup> A collection of recent attempts to resolve Benacerraf's dilemma in these ways is to be found in Irvine (1990).

recognizing the common ground of mathematics and other human cognitive enterprises, i.e. there is a growing number of people who consider the similarities and analogies which mathematics and the empirical sciences share, rather than the previously overestimated difference.

5. Hence, very recently Maddy (1991) announced the "post-Benacerrafian-age" of philosophy of mathematics. According to her, Benacerraf's dilemma no longer has the determining character it had had for such a long time, and the post-Benacerrafian philosopher of mathematics has to look for a new agenda. What should it be? According to Maddy we are left with two main unsettled questions:

- (1) Which of the various ontological tinkerings proposed to dissolve Benacerraf's dilemma is best?
- (2) How can the axioms of set theory be justified?

According to Maddy, the first question is a parochial squabble between philosophers. The other question, however, "is a philosophical inquiry that could provide guidance for actual mathematics where guidance is sorely needed" (Maddy 1991:158). We think that in some sense Maddy's first assessment is correct. Just as in the case of the empirical science the ontological details are not as important as philosophers traditionally used to make us believe: modern sciences do not take ontological questions too seriously - after all, do we know what an electron "really" is, and does this question make sense at all? In any case, it is not a central question of contemporary philosophy of science. Notwithstanding the basic underlying assumption of classical philosophy of mathematics, mathematics too might exhibit this kind of ontological indifference or indeterminacy. Several contributions to this volume deal with this phenomenon: Mac Lane emphasizes the *protean* character of mathematics, i.e. the fact that one and the same mathematical structure has many different realizations, and it does not make sense to ask "which is the essential one?". In a similar vein Grattan-Guinness and Resnik stress the fact that mathematics comprises a great diversity of forms, reasonings, structures, and applications. Hence there is no all embracing single answer to the questions such as "What is it that a mathematical theory talks about?" and "How is a mathematical theory applied to a domain of empirical phenomena?".

Thus a kind of branching of philosophy of mathematics in two disciplines seems to take place (it might even be that it has already occurred): following Grattan-Guinness the first may be dubbed *philosophers' philosophy of mathematics*, dealing exclusively with logics, set theory, and perhaps elementary arithmetics of natural numbers, whereas the other one may be called

*mathematicians' philosophy of mathematics* where *real mathematics* is taken into account to a greater extent. We are sure that such a more "extroverted" philosophy of mathematics is urgently needed, and philosophy of mathematics should follow Mac Lane's admonition "that the philosophy of mathematics should be based on observation of what is actually present in the subject and not, as so often, on a few speculative notions about the most elementary parts of mathematics". Taking this seriously implies that mathematics is to be seen *not* as a single theory but as an extended network of theories related to each other through intertheoretical relations of various kinds. The task to provide a typology of the elements and the relations which form this network, and the further question how it develops, are surely important subjects for philosophy of science and mathematics. The contributions of Moulines on intertheoretical relations, and of Sneed on machine implemented discovery of scientific theories, deal with these topics.

Thus, mathematics provides a huge space of concepts and phenomena that cannot be surveyed from the perspective of an armchair philosopher having just some elementary fragments of school mathematics within easy reach. This amounts to a certain naturalism in philosophy of mathematics: The idea that philosophy could provide a justification or foundation for mathematical method has to be abandoned. To a large extent, the foundations of mathematics are provided by mathematics, and the role of philosophy as a founding discipline for mathematics seems to be limited. This, of course, is a controversial thesis that needs further elaboration. In particular, it depends on how the twins of "introverted" philosophy of mathematics concerned with traditional topics of philosophy of mathematics such as foundations, existence of mathematical objects, etc. and "extroverted" philosophy of mathematics concerned more directly with the reality of the discipline can be related. In any case it does not seem desirable that both should live totally separated from each other. The desideratum of a truly *general philosophy of mathematics* (Grattan-Guinness) is still to be met.

What about Maddy's second proposal, modern philosophy of mathematics should primarily deal with the question of how the set theoretical axioms could be justified? In our opinion this assertion needs modification<sup>5</sup>. The

---

<sup>5</sup> Her reason for concentrating on the question of set theoretical axioms is telling: "For foundational purpose, the only axioms of current mathematics appear in set theory. The characteristic objects of other branches of mathematics are officially defined within set theory" (Maddy 1991:158). We don't agree with any of these assumptions: as is argued in many of the essays in this volume neither the philosophy of mathematics should be restricted to

concentration on set theoretical axioms offers a far too narrow perspective on what mathematical knowledge is and how it develops. We believe the contributions of the present volume give a lot of useful clues on how a more comprehensive agenda for a "New" philosophy of mathematics might look. One topic surely will be the multiple and variegated relations of mathematics with empirical science. In this task a natural ally for the philosopher of mathematics is, or at least should be, his colleague from the department of philosophy of science.

6. What has modern philosophy of science to tell us about mathematics? Philosophy of science has been philosophy of the *empirical* sciences. Hence, any answer to this question is driving at the similarities or analogies between mathematical and empirical knowledge. Quite different approaches may be distinguished. One might focus on empirical traces inside mathematics. The by now classic example of this approach is Lakatos' "quasi-empiricist" philosophy of mathematics. Perhaps the most "Lakatosian" approach in this sense is provided by Echeverría who tries to show the quasi-empirical character of number theory. Niiniluoto critically evaluates some of the main ideas of the "quasi-empiricist" approach relating it with a Popperian ontology. Breger points out that even for mathematics tacit knowledge, i.e. not directly formalizable knowledge about mathematics plays an essential role in the historical development of the discipline.

Another possibility to tap our understanding of the empirical sciences for a modern philosophy of science for the philosophy of mathematics is to study similarities, analogies, and relations of mathematics and empirical science. In different ways this is performed in the contributions of Ibarra/Mormann (Structural analogies between mathematical and empirical science), Rantala (Reduction in mathematics and empirical science), Resnik (Indispensability of mathematics for science) and Torretti (Mathematical versus physical necessity), and Schmidt (Empirical meaning of set theoretical axioms), Scheibe (Non-conservative embedding of physical theories in mathematics), Da Costa/Doria (Metamathematical phenomena within mathematical physics) and Mosterín (Mathematical description as an encoding process) concentrate on different dimensions of the problem of applying mathematics in science.

However, the relation of modern philosophy of science and philosophy of mathematics is not a one-way street. It also makes sense to ask "What has modern mathematics to tell philosophy about science?" It might well be the

---

the task of foundations, nor can set theory be considered as the overarching frame of all mathematics.

case that the long lasting preoccupation of philosophy of science with empirical sciences might have blinded it to some important aspects not only of mathematics but also of the empirical sciences themselves. In our opinion, an important example might be provided by the fact that the development of science should be understood not as a (cumulative) piling up of more and more theorems but rather as a conceptual development. Quite a few contributions in this volume, from the philosophical as well from the historical strand endorse this point of view: Mac Lane and Lawvere point out that understanding a piece of mathematics for the individual as well as for the scientific community of mathematicians as a whole is a long, and probably never ending process. Lawvere deals with the dialectical development of the categories of Space and of Quantity, and in detailed historical studies Feferman and Mahoney study the conceptual development of recursion theory and the mathematics of computing. Thus, an important contribution philosophy of mathematics could offer to philosophy of science might be the insight into the importance of the conceptual dimension of scientific knowledge.

The communication between historians, philosophers, and mathematicians on a multifaceted and complex topic such as mathematics cannot always be easy. We hope, however, that the present collection of essays shows that it can be carried out in a fruitful way leading to new promising explorations in the fascinating space of mathematics.

The Editors

### References

- Aspray, W., Kitcher, Ph.: 1988, "An opinionated Introduction", in W. Aspray, Ph. Kitcher, *History and Philosophy of Modern Mathematics*, Minnesota Studies in the Philosophy of Science XI, Minneapolis, University of Minnesota Press, 3-57.
- Hersh, R. (ed.): 1991, "New Directions in the Philosophy of Mathematics", *Synthese* 88(2).
- Irvine, A.D. (ed.): 1990, *Physicalism in Mathematics*. The University of Western Ontario Series in Philosophy of Science, vol. 45, Dordrecht, Kluwer.
- Maddy, P.: 1991, "Philosophy of Mathematics: Prospects for the 1990s", *Synthese* 98, 155-164.
- Tymoczko, Th. (ed.): 1985, *New Directions in the Philosophy of Mathematics*, Boston, Birkhäuser.

## Structural Dimensions



# The Protean Character of Mathematics

SAUNDERS MAC LANE (Chicago)

## 1. Introduction

The thesis of this paper is that mathematics is protean. This means that one and the same mathematical structure has many different empirical realizations. Thus, mathematics provides common overarching forms, each of which can and does serve to describe different aspects of the external world. This places mathematics in relation to the other parts of science: mathematics is that part of science which applies in more than one empirical context.

This paper will first present some of the evidence for this thesis. Much of this evidence is essentially common knowledge, as we will illustrate. Additional such evidence has been presented in my book *Mathematics: Form and Function*,<sup>1</sup> Springer Verlag, 1985. There and here, we follow the important, but often neglected principle that the philosophy of mathematics should be based on observation of what is actually present in the subject and not, as so often, on a few speculative notions about the most elementary parts of mathematics.

Finally, we will draw a number of consequences from our thesis.

## 2. Arithmetic is Protean

At the very beginning of our subject, observe that the natural numbers have more than one meaning. Such a number can be an ordinal: first, second, or third.... Or it can be a cardinal: one thing, two things,.... The natural number two is thus neither an ordinal nor a cardinal; it is the number two, with these two different meanings to start with. It is the form of "two", which fits different uses, according to our intent. As a result, the formal introduction of these natural numbers can be made in different ways – in terms of the Peano postulates (which describe not unique numbers, but the properties which such num-

---

<sup>1</sup> Berlin-New York, Springer Verlag, 1985.



bers must have) or in terms of cardinals – two is the set of all sets of unordered pairs – or in terms of ordinals, where two is the set of all ordered pairs, and so on. This and other definitions have alternative forms, as with the von Neumann description of the ordinals, in which 3, for example, is the set  $\{0, 1, 2\}$  of all smaller ordinals.

Hence, at the very beginning of mathematics, natural numbers are not objects, but forms, variously described with a view to their various practical meanings. Put differently, an axiomatic description of number, as with Peano, does not define THE NUMBERS, but only numbers up to isomorphism. This recognition of the prevalence of mathematical descriptions "up to isomorphism" has recently been reemphasized in category theory, where products, adjoints and all that are inevitably defined only "up to an isomorphism".

Numbers are little without the corresponding arithmetic operations, addition, multiplication... But addition has different explanations; for cardinals, addition is the disjoint union of the corresponding sets; for ordinals addition is given in terms of suitable successors, as in the familiar definition of addition by recursion:

$$m + 0 = m, \quad m + (n + 1) = (m + n) + 1.$$

Multiplication has even more different introductions, as repeated addition, or by a recursion, as in

$$m \cdot 0 = 0, \quad m(n + 1) = mn + m,$$

or by sets of ordered pairs, or whatever. In line with different definitions of addition, there are even more different uses: multiplication to calculate area (width times height) or to calculate cost (price times number) or the number of inches (12 times number of feet). Multiplication is important not because it has a fixed meaning but because there are many different applications of the one idea of a "product".

Laws of arithmetic are similarly of varied meaning. One word, "associativity" covers both a property of addition

$$m + (n + k) = (m + n) + k$$

and a property of multiplication

$$x(yz) = (xy)z.$$

The same formal property of a binary operation, under the same name, comes up elsewhere in algebra – for the tensor product of vector spaces, for the intersection numbers in algebraic geometry, and for the product of cocycles in algebraic topology. The associative law appears as an axiom in the definition of a

group, or of a semigroup, or of a Hopf algebra. For all these and other varied uses there is really just the one form, "associative", present whatever the name of the operation involved – addition, multiplication, whatever. The content varies greatly, but not the form.

### *3. Geometry is Protean*

Just as for arithmetic, geometry is protean – the same figure in many different guises, the same idea described in many different ways. Thus similarity: Two triangles are similar when corresponding angles are equal, or when the ratios of corresponding sides are the same, or when one triangle can be obtained from the other by translation, followed perhaps by rotation or reflection. These definitions are demonstrably equivalent ways of formulating one meaning. But the meaning is after all not restricted to triangles: The definition in terms of motions of translation, etc. applies to the similarity of more complex figures in the plane or in space. "Similarity" is a protean idea.

The subject matter of geometry enjoys a like ambiguity. The formal study of geometry begins usually with plane geometry. But which plane? That assumed by Euclid? Perhaps not, for his axioms were after all not quite complete. They were firmed up in the famous axiomatics by Hilbert, which filled the gap left by the absence of the Pasch axiom, but is Hilbert's plane the real one, or do we really think of the plane in terms of coordinates, where the cartesian plane has points determined by ordered pairs  $(x, y)$  of numbers  $x$  and  $y$ ? That does indeed give a plane, but those numbers do depend upon a choice of origin and of axes through that origin – and that choice has little to do with the Plane "an sich". And after all if the given is the three dimensional world, the plane is an abstraction – which plane in that three dimensional continuum do you mean? A question without a good answer; no matter, because the plane is a form which has many different realizations. For mathematics, one contemplates a "plane", because it can be studied profitably by itself, and then used in its many and various incarnations.

### *4. Analysis*

To advance to higher reaches: Calculus is protean. Take for example the derivative  $dy/dx$  of a function  $y$ . If  $y$  and  $x$  are coordinates, the derivative means the slope of the corresponding curve. If  $y$  is distance and  $x$  is time, the derivative is velocity, but if  $y$  itself is velocity, the derivative is acceleration. If  $y$  is total cost and  $x$  the number purchased, the derivative is a marginal cost, and this same idea of the marginal occurs elsewhere in economics. Perhaps one

might say generally that  $dy/dx$  is a rate of change, but this seems to imply that  $x$  is something like a time, and this is not always so. Formally, one may define the derivative as the limit of the ratio of increments in  $x$  and  $y$ , but this requires epsilons and deltas and may lose sight of some of the preceding meanings. Thus the derivative is really a general form, with many different interpretations.

Calculus, in its second parallel aspect, deals with integrals. The area under a curve is an integral. So is a volume. So is a moment of inertia, or the pressure of water on a dam, or any one of the many other examples with which beginning students of the calculus are tormented. Or again, integration is the reverse of differentiation. Which is the real integral? In some sense, all are; in a better sense, the integral is a general notion with many different representations in mathematics and in the world.

For differential equations, much the same is true. One learns how to solve this or that ordinary differential equation – because it often appears in different applications. And the same partial differential equation can apply to the motion of water, or of air, or of heat, or of light.... Thus throughout calculus (and the higher analysis as well) one and the same idea turns out to fit many different sets of facts – a fortunate conclusion for the efficient progress of science.

Put differently, Isaac Newton may have been stimulated by a falling apple – but his work was not limited to falling apples or to other falling objects, or to planets or to tides or to economic changes. The mathematics matters because it is present in many different physical forms.

### *5. The Unexpected*

Many connections of mathematical ideas are unexpected; a mathematical form, studied for one purpose later crops up in a different context. Thus tensor analysis, with its initial welter of subscripts, was first built to handle higher curvatures and transportation of structures in differential geometry. Then, suddenly, it was used by Einstein to formulate general relativity in terms of his tensor equations. Then later, it turned out that those confusing multiple subscripts on the tensors were not necessary, since a tensor could be described conceptually as an element in a suitable "tensor product" space. That view arose because those same tensor products were needed for cohomology in algebraic topology. More recently, they crop up as typical adjunctions in the comparison of different categories and topoi. The idea of "tensor", once formulated, spreads far and wide.

A group is an algebraic object – a set of elements any two of which can be

multiplied so that the product  $xy$  of elements is associative – and has a unit and an inverse. That is a formal definition, but the form was extracted from cases of symmetry-symmetry in the solution of algebraic equations and in the description of the varieties of crystal structure or of ornamental symmetry. Once groups were abstracted, mathematicians turned to the description of concrete representations of groups, say by groups of transformations (or matrices). With the initial work of Frobenius and I. Schur, this was a piece of presumptively pure mathematics. But then, with Hermann Weyl and others, group representations cropped up in quantum mechanics, at first to be labelled as a "Gruppenpest" and then gradually to be accepted as the right way – sometimes as the "eight fold way". A group is an abstract object, as are its representations – but they appear in many different real contexts.

There are many other, more limited examples. Thus early in this century, the Austrian mathematician Radon studied the way in which the values of a three dimensional integral could be reconstituted by suitable two-dimensional integrals of the underlying quantities. This must have seemed a pure mathematical exercise. Now with the development of a medical tool, the Cat-scan, it has all manner of uses to reconstruct, say, a three dimensional image of a patient's brain from two-dimensional x-ray pictures. Radon did not set out to study the brain, but that study used – and developed further – the forms of analysis which he had developed.

In the nineteenth century, Riemann studied the singularities of functions of a complex variable, and found a partial formula for the number of such functions with a given array of singularities. Improved, this is now given by the Riemann-Roch formula. For the behavior of algebraic curves that formula has geometric meaning, but it underwent a vast geometric generalization at the hands of Hirzebruch and Grothendieck, giving rise to a special branch of topology known as K-theory. Still more recently it appears as part of the Atiyah-Singer index theorem for differential operators – and recently, ideas from mathematical physics have produced better proofs for this index theorem.

Theoretical physics yields many more examples of the unexpected. I now cite one in which I happen to have been involved. In 1963, Stasheff and I studied a case of associativity – where the associative law is not an identity, but just an isomorphism as in the operation, left to right:

$$a: \quad x (y z) \rightarrow (x y) z.$$

This operation moves parentheses back to front. If it is the correct operation, it must satisfy a condition for four factors, by moving parentheses back to front in two ways, as in

$$x (y (z t)) \rightarrow ((x y) z) t.$$

For this, it turned out that  $a$  must satisfy a certain pentagonal condition, using the two different ways of moving parentheses back to front in this formula. Then if also the commutative law entered, there was a hexagonal condition involved. Suddenly, in 1988, a pentagon and a hexagon turned up in some physics studies of "conformal field theory". It was the same pentagon and hexagon. In this way what had seemed an abstract algebraic example of associativity and commutativity had a different contact with reality. Unexpected, but not unreasonable, since the group representations mentioned above are really involved here, as well as some traditional questions about knots and braids.

Theoretical computer science provides many more examples of the unexpected appearance of mathematical concepts in new context. The most striking case is that of the lambda calculus. This had been first formulated about 1930 by Alonzo Church as a possible new foundation of mathematics. Its basis was the lambda operation, which turns an expression  $f(x)$  into the function  $f$ , or more extensively, a expression  $f(x,y)$  with two variables into the corresponding  $f(-,y)$ : that function of  $x$  yielding a function of the remaining variable  $y$ . This is just the process familiar in the calculus, whereby a partial derivative is reduced to an ordinary one. As a foundation of mathematics, this lambda calculus did not succeed, but the formalism itself lived and presently was the inspiration for the programming language "LISP". This was just the first step in the current massive usage of seemingly technical ideas from mathematical logic in computer science, as in the way proof theory is used to describe data types, and in the extensive other uses of type theory from logic. This came at a time when type theory, in its classical Russellian form, had almost become wholly obsolete. And the rapid changes of fashion in computer science concern not only logic, but questions about algorithms, levels of computability, combinatorics, and the use of categories to formulate properties of polymorphic data types.

### *The Consequences of the Protean Character of Mathematics*

#### *6. The Formal*

As long ago observed, mathematics is intrinsically formal. This is a consequence of its protean character: because mathematics is not about this or that actual thing, but about a pattern of form suggested by various things or by previous patterns. Therefore mathematical study is not study of the thing, but of the pattern – and thus is intrinsically formal. Properties of things may suggest theorems or provide data, but the resulting mathematics stands there independent of these earlier suggestions. The actual mathematics proceeds by following rules of calculation or of algebraic manipulations or by use of

collections of already established formulas or by calculations on a computer or by logical deductions according to accepted rules of inference. This is the reason that, in the long run, mathematics is essentially axiomatic. It is true by reference to the axioms and not to the facts.

### *7. Foundations*

It follows that mathematics does not need a "Foundation". Any proposed foundation purports to say that mathematics is about this or that fundamental thing. But mathematics is not about things but about form. In particular, mathematics is not about sets. Zermelo and then Zermelo-Fraenkel (and Skolem) did formulate axioms for sets. They are at times convenient ways of coding pieces of mathematics, but they do not catch the reality. For example, at the very beginning, the set-theorist's description of the ordered pair  $\langle x, y \rangle$  as the set  $\{x, \{x, y\}\}$  is clearly an artificial arrangement. Von Neumann's ordinal numbers are a convenience and not a reality. A real number is not a Dedekind cut nor a Cantor sequence. Real numbers live in mathematics precisely because of their multiple meanings. No one meaning is "it".

It is a curious historical observation that the popularity of a set theoretic foundation came with Bourbaki and then with the new math just about the time when the uniqueness of the set theoretic foundation was challenged by Lawvere with alternative categorical axiomatizations.

Thus we must conclude that there just does not exist a platonic world of sets, that large cardinals live in a never-never land, and that the continuum hypothesis will remain forever unsettled – Gödel proved it consistent with the Zermelo-Fraenkel axioms, while Paul Cohen proved that it is independent of these axioms. Current attempts to settle the continuum hypothesis are therefore futile. Understanding the nature of mathematics is an effective guide to productive directions of research.

Because of its protean nature, mathematics does not need a foundation – not by Plato, or Frege, or Whitehead-Russell, or Zermelo, or Hilbert, or Brouwer, or Quine, or Bishop. This is because it does not deal with one ultimate substance but with the forms common to different substances. Instead what mathematics does is to codify varied rules for calculation and for proofs and to organize and deepen understanding of what has been done.

### *8. Dead Ends*

Good understanding of the nature of mathematics helps us to realize when an apparent part of mathematics is in fact a dead end. One example is the notion

of a "fuzzy" set. In an ordinary set  $S$ , one knows of each potential element  $x$  only whether it is in, or not in,  $S$ . For a fuzzy set, one has instead a "level" of membership: a function  $f(x)$ , usually ranging from 0 to 1, specifying the extent to which  $x$  belongs to the intended fuzzy set. At their origin such fuzzy sets were used to describe certain engineering control devices – things partly off or partly on according to circumstance. Thus they may have limited applications. But then the doctrine that mathematics is about sets took hold, and various people started to rewrite this or that part of mathematics with "set" replaced everywhere by "fuzzy set". This happened, for example, with topology, where fuzzy topology spaces grew rampant without any real relation to geometry or topology, but only with the aim of making general topology fuzzier. The result was more paper and not more progress.

There are many other examples of mathematical dead ends. For instance, in universal algebra a so-called "groupoid" is a set with one binary operation, no further properties required – and hence no properties of consequence. Similar excesses have arisen in graph theory. A famous old theorem of Kuratowski specifies when a graph cannot be drawn in the plane without unwanted crossing: it must then contain either the complete graph on five points or a graph on two sets of three points with each point in the first set joined by an edge to each in the second. This Kuratowski theorem was a striking one. Unhappily, specialists tried to do the same sort of thing for graphs which cannot be drawn on the torus. They found a vast – and hence totally uninteresting – list of such "non-toral" graphs. In number theory, there is the fundamental theorem that every whole number can be written uniquely as a product of prime numbers. Nowadays, enthusiasts with big computers contend with each other as to who can so factor the biggest yet number. This becomes a display of virtuosity, hardly justified by the claimed use of factoring in breaking codes. Similar troubles arise in many applications of mathematics. Gauss invented least square approximations, and this can be used to make regressions – getting the best linear fit for given data in many variables. There are now canned programs which will do all the routine work of such fitting – as a result things are fitted and "shadow" costs are determined from data that is incomplete and misleading, often because only some of the relevant variables appear.

Here and elsewhere good mathematics requires discrimination. We need judgement and not inappropriate foundations or excess calculations.

### *9. Interrelated Networks*

In mathematics we have seen that the same form may represent different facts

– the same symmetric group represents the permutations of the roots of an algebraic equation or the interchanges among the atoms of a molecule or the symmetries of a tetrahedron in higher dimensions. Or, the continuum of the reals is either axiomatized geometry in one dimensions or an extension of the rational number system or the basic range for physical measurements or...

From this we conclude that mathematics is actually a tightly connected network of different forms and concepts. It combines rules, formulas, formal systems, theorems, applications, concepts, and algorithms, all of them extracted over the years from the facts of the world. One can then write extensive network diagrams – as in my cited book – to display how these mathematical centers are intertwined. One might say that the subject begins in the human experiences of:

MOVING MEASURING SHAPING COMBINING COUNTING,

and that these lead, more or less in that order, to disciplines such as

APPLIED MATHEMATICS CALCULUS GEOMETRY ALGEBRA NUMBER THEORY.

### *10. Progress*

Advances in the multiple realms of mathematics involves two complementary things: the solution of outstanding problems and the understanding of achieved results.

There are many famous old problems in mathematics. There is Fermat's last theorem: to find two whole numbers whose  $n$ th powers for  $n > 2$  add up to a single  $n$ th power. There is the Riemann hypothesis: where are the zeros of the zeta function – a function which controls much of the distribution of the prime numbers among the other integers. There is the Poincaré conjecture: to characterize the three dimensional sphere by homotopy properties from among other three dimensional manifolds (It has been solved in dimensions higher than 3, most recently in the difficult case of four dimensions, but the original case of the 3-sphere remains still unsettled).

From time to time, some of these problems are solved. Thus, the Mordell conjecture stated that most polynomial equations, like that of Fermat, have at most a finite number of solutions in whole numbers. This has recently been settled by Faltings, who needed to use for this purpose some of the recent extensive abstract techniques of algebraic geometry. More unsolved problems await us.

The other goal of mathematics is understanding. For example, a century ago, three dimensional space, where points are determined by three coordinates,



was reduced to a space of vectors. At first this seemed to be just a technique to write one vector equation instead of three equations in coordinates – but it is actually much more. Vector spaces visualize the solutions of simultaneous linear equations, they formulate electrodynamic fields and other fields in physics, they allow geometry in more dimensions, even in infinite ones, and they provide a setting for the functional analysis of linear operators in analysis.

There are many other ways in which greater understanding develops. Matrix multiplication is understood not by multiple subscripts, but by multilinearity of the tensor product, and so on and on.

### 11. An Example

As an example of the long process of understanding a piece of mathematics, I take the case of Galois theory, dealing with the solution of algebraic equations. There, each quadratic equation can be solved by a familiar formula involving a square root – but it turns out that, for an equation of degree 5 or higher, no such solution by roots ("radicals") is possible. The reason was finally seen to lie in the study of the symmetries of an equation – the group of allowed permutations of its roots. Here is a list of some of the people who have contributed to the real understanding of this situation.

Lagrange, 1770	He found certain resolvents
Abel, 1827	Degree 5 by radicals is impossible
Galois, 1830	Groups and subgroups enter decisively
Jordan, 1870	Wrote his <i>Traité des substitutions</i> on groups
Dedekind, 1900	Glimpsed the conceptual formulation
Steinitz, 1910	Included abstract fields in this study
Emmy Noether, 1922	Saw the power of the conceptual approach
van der Waerden, 1931	Wrote it up in <i>Moderne Algebra</i>
J. F. Ritt, 1932	Extended Galois theory to differential algebra
Emil Artin, 1938	Brought in properties of linear dependence
Birkhoff/Mac Lane, 1941	Presented this in <i>Survey of Modern Algebra</i>
Jacobson, 1944	Included inseparable extensions of fields
Kolchin, 1946	Brought in differential Picard-Vessiot theory
Kan, 1957	Adjunctions explain Galois correspondences
Grothendieck, 1961	Introduced Galois theory for covering spaces
Chase-Rosenberg, 1965	Galois theory for rings (with Harrison)
Kaplansky, 1969	Lucid presentation, with examples and problems
Joyal/Tierney, 1981	Generalized Grothendieck, with locales
Kennison, 1983	Theaters of action for Galois theory

Janelidze, 1988

Categorical formulation of Galois structure

Various, 1990

One adjunction handles Galois and much more.

This example of growing understanding could be supplemented by many other cases of the gradual understanding of the protean and interconnected forms of mathematics. May protean understanding prosper!

## Categories of Space and of Quantity

F. WILLIAM LAWVERE (New York)

0. The ancient and honorable role of philosophy as a servant to the learning, development and use of scientific knowledge, though sadly underdeveloped since Grassmann, has been re-emerging from within the particular science of mathematics due to the latter's internal need; making this relationship more explicit (as well as further investigating the reasons for the decline) will, it is hoped, help to germinate the seeds of a brighter future for philosophy as well as help to guide the much wider learning of mathematics and hence of all the sciences.
1. The unity of interacting opposites "space vs. quantity", with the accompanying "general vs. particular" and the resulting division of variable quantity into the interacting opposites "extensive vs. intensive", is susceptible, with the aid of categories, functors, and natural transformations, of a formulation which is on the one hand precise enough to admit proved theorems and considerable technical development and yet is on the other hand general enough to admit incorporation of almost any specialized hypothesis. Readers armed with the mathematical definitions of basic category theory should be able to translate the discussion in this section into symbols and diagrams for calculations.
2. The role of space as an arena for quantitative "becoming" underlies the qualitative transformation of a spatial category into a homotopy category, on which extensive and intensive quantities reappear as homology and cohomology.
3. The understanding of an object in a spatial category can be approached through definite Moore-Postnikov levels; each of these levels constitutes a mathematically precise "unity and identity of opposites", and their ensemble bears features strongly reminiscent of Hegel's *Science of Logic*. This resemblance suggests many mathematical and philosophical problems which now seem susceptible of exact solution.

*0. Renewed Progress in Philosophy Made Both Necessary and Possible by the Advance of Mathematics*

In his Lyceum, Aristotle *used* philosophy to lend clarity, directedness, and unity to the investigation and study of particular sciences. The programs of Bacon and Leibniz and the important effort of Hegel continued this trend. One of the clearest applications of this outlook to mathematics is to be found in the neglected 1844 introduction by Grassmann to his theory of extensive quantities. Optimistic affirmations and applications of it are also to be found in Maxwell's 1871 program for the classification of physical quantities and in Heaviside's 1887 struggle for the proper role of theory in the practice of long-distance telephone-line construction. In the latter, Heaviside formulates what has also been my own attitude for the past thirty years: the fact that our knowledge will of course never be complete, and hence no general theory will be final, is no excuse for not using now the most general theory which science can support, and indeed for accuracy we must do so.

To students whose quest drives them in the above direction, the official bourgeois philosophy of the 20th century presents a near vacuum. This vacuum is the result of the Jamesian trend clearly analyzed by Lenin in 1908, but "popularized" by Carus, Mauthner, Dewey, Mussolini, Goebbels, etc. in order to create the current standard of truth in journalism and history; this trend led many philosophers to preoccupation with the flavors of the permutations of the thesis that no knowledge is actually possible. Naturally this 20th century vacuum has in particular tried to suck what it can of the soul of mathematics: a science student naively enrolling in a course styled "Foundations of Mathematics" is more likely to receive sermons about unknowability, based on some elementary abstract considerations about subjective infinity, than to receive the needed philosophical guide to a systematic understanding of the concrete richness of pure and applied mathematics as it has been and will be developed.

By contrast, mathematics in this century has not been at a standstill. As a result mathematicians at their work benches have been forced to fashion philosophical tools (along with those proofs of theorems which are allegedly their sole product), and to act as their own "Aristotles" and "Hegels" as they struggle with the dialectics of 'general' and 'particular' within their field. This is done in almost complete ignorance of dialectical materialism and often with understandable disdain for philosophy in general. It was struggle with a problem involving spheres and the relation between passage to the limit and the leap from quantity to quality which led Eilenberg and Mac Lane in the early 1940's to formulate the general mathematical theory of categories, functors, and natural transformations. Similarly, study of concrete problems in algebraic

topology, functional analysis, complex analysis, and algebraic geometry in the 1950's led Kan and Grothendieck to formulate and use important further advances such as adjoint functors and abelian categories. And the past thirty years have not been devoid of progress: from the first international meeting on category theory in La Jolla, California in 1965 to the most recent one in Como, Italy in 1990, toposes, enriched categories, 2-categories, monads, parameterized categories (sometimes called "indexed"), synthetic differential geometry, simplicial homotopy, etc. have been refined and developed by over two hundred researchers with strong ties to nearly every area of mathematics. In particular all the now-traditional areas of subjective logic have been incorporated with improvement into this emerging system of objective logic.

It is my belief that in the next decade and in the next century the technical advances forged by category theorists will be of value to dialectical philosophy, lending precise form with disputable mathematical models to ancient philosophical distinctions such as general vs. particular, objective vs. subjective, being vs. becoming, space vs. quantity, equality vs. difference, quantitative vs. qualitative etc. In turn the explicit attention by mathematicians to such philosophical questions is necessary to achieve the goal of making mathematics (and hence other sciences) more widely learnable and useable. Of course this will require that philosophers learn mathematics and that mathematicians learn philosophy. I can recall, for example, how my failure to learn the philosophical meanings of "form, substance, concept, organization" led to misinterpretation by readers of my 1964 paper on the category of sets and of my 1968 paper on adjointness in foundations; a more profound study of Hegel's *Wissenschaft der Logik* and of Grassmann's *Ausdehnungslehre* may suggest simplifications and qualitative improvements in the circle of ideas sketched below.

### *1. Distributive and Linear Categories; The Functoriality of Extensive and Intensive Quantities*

A great many mathematical categories have both finite products and finite coproducts. (A product of an empty family is also known as a terminal object, and an empty coproduct as a coterminal or initial object). However, there are two special classes of categories defined by the validity of two special (mutually exclusive) relationships between product and coproduct. One of these may be called *distributive* categories, for these are defined by the requirement that the usual *distributive law* of arithmetic and algebra should hold for multiplication (=product) and addition (=coproduct) of objects, in the precise sense that the natural map from the appropriate sum of products to a product of sums should be an isomorphism; this includes as a special case that

the product of any object by zero (=initial object) is zero. The other class of *linear* categories is defined by the requirement that products and coproducts *coincide*; more precisely, a coterminal object is also terminal in a linear category, which permits the definition of a natural map (= "identity matrix") from the coproduct of any two objects to their product, and moreover this natural map is required to be an isomorphism. As pointed out by Mac Lane in 1950, in any linear category there is a unique commutative and associative addition operation on the *maps* with given domain and given codomain, and the composition operation distributes over this addition; thus linear categories are the general contexts in which the basic formalism of linear algebra can be interpreted.

All toposes are distributive. General categories of discrete sets, of continuous sets, of differentiable, measurable, or combinatorial spaces tend to be distributive, as do categories of non-linear dynamical systems. Given a particular space, there are categories of sheaves on it, of covering spaces of it, etc. which provide an expanded or restricted view of what happens in that particular space and are also distributive. Since both general ("gros") and particular ("petit") spatial categories are distributive categories, a useful philosophical determination would be the identification of "categories of space" with distributive categories. Since distributive categories such as that of the permutation representations of a group can often be seen to be isomorphic with spatial categories such as that of the covering spaces of a particular space having that group as fundamental group, the inverse identification has merit; it also permits to use geometrical methods to analyze categories of concepts or categories of recursive sets. For many purposes it is useful to "normalize" distributive categories by replacing them with the toposes they generate, permitting application of the higher-order internal logic of topos theory to the given distributive category; on the other hand many distributive categories are "smaller" than toposes and in particular have manageable Burnside rigs. Here by "rig" we mean a structure like a commutative ring except that it need not have negatives, and the name of Burnside was suggested by Dress to denote the process of abstraction (exploited recently by Schanuel) which Cantor learned from Steiner: the isomorphism classes of objects from a given distributive category form a rig when multiplied and added using product and coproduct; the algebra of this Burnside rig partly reflects the properties of the category and also partly measures the spaces in it in a way which (as suggested by Mayberry) gives deeper significance to the statement attributed to Pythagoras: "Each thing is number". Still in need of further clarification is the contrast within the class of distributive categories between the "gros" (general category of spaces of a certain kind) and the "petit" (category of variable sets over a

particular space); this distinction (a qualitative one, not one of size) has been illuminated by Grothendieck and Dubuc and, I hope, by my 1986 Bogotá paper [14]; these show the importance of the ways in which an object in a "gros" category can give rise to a "petit" category, and the additional "structure sheaf" in the "petit" category which reflects its origin in the "gros" environment.

The category of real "vector spaces", the category of abelian groups, the category of topological vector spaces and the category of bornological vector spaces are all linear categories. So are the category of projective modules over any particular rig and the category of vector bundles over any particular space. In the last example, the vector bundles (=objects) themselves are *kinds* of variable quantities over the space, and the *maps* between these are particular variable quantities over the space. Thus "categories of quantity" will be tentatively identified with linear categories. Abelian AB5 categories are special linear categories having further "exactness" properties; again "normalization" may be useful, even within functional analysis. For abelian categories and many others, the Mac Lane addition of maps is actually an abelian group, that is, each map has a negative. However, for some other linear categories addition is actually idempotent (and hence could not have negatives in this algebraic sense); this occurs in logic (in the narrow sense) where the quantities are variable truth values (reflecting "relations"), and in geometry when quantities are (variable) dimensions and the multiplication is *not* idempotent.

What is a space and how can quantities vary over a space? We have suggested above that, formally, a space is either a "petit" distributive category or an object in a "gros" distributive category. But as spaces actually arise and are used in mathematical science, they have two main general conceptual features: first they serve as an arena for "becoming" (there are spaces of states as well as spaces of locations) and secondly they serve as domains for variable quantity. These two aspects of space need to be expressed in as general a mathematical form as possible: in section 2, I will return to "becoming" and one of its roles in mathematics, but in this section I concentrate on the relation between space and variable quantity.

Broadly speaking there are two kinds of variable quantity, the extensive and the intensive. Again speaking broadly, the extensive quantities are "quantity of space" and the intensive quantities are "ratios" between extensive ones. For example, mass and volume are extensive (measures), while density is intensive (function). Although Maxwell managed to get extensive quantities accepted within the particular science of thermodynamics, and although Grassmann demonstrated their importance in geometry, there is still a reluctance to give them status equal to that of functions and differential forms; in particular the use of the absurd terminology "generalized function" for such distributions as

the derivative of the Dirac measure has created a lot of confusion, for as Courant in effect observed, they are not intensive quantities, generalized or not. "Generalized measure" would have been a better description of distributions; to show that a distribution "is a function" involves finding a density for it relative to a "fixed" reference measure, but only in special non-invariant circumstances do the latter exist.

Broadly, a "type of extensive quantity" is a covariant coproduct-preserving functor from a distributive category to a linear category. The last condition reflects the idea that if a space is a sum of two smaller spaces, then a distribution of the given type on it should be determined by an arbitrary pair of distributions, one on each of the smaller spaces, while by the defining property of a linear category, "pairs" are equally well expressed in terms of coproducts in the codomain of our functor. The covariant functoriality has itself non-trivial consequences: the value of the functor at the terminal space may be considered to consist of *constant* quantities of the given type, and the value of the functor at a given space to consist of the extensive quantities of the given type which are *variable* over that space; since any given space has a unique map to the terminal space, the functor induces a map in the linear category which assigns to each variable extensive quantity its *total*, which is a constant. For example, the quantity of smoke now in my room varies over the room, but in particular has a *total*. On the other hand a map *from* the terminal space *to* a given space is a *point* of that space; thus the functor assigns to such a point a linear map which to any constant weight of the given type assigns the Dirac measure of that weight which is supported on that point. For a more particular example of the covariant functoriality in which neither domain nor codomain of the inducing map reduces to the terminal space, consider the following definition of the term *sojourn*: the extensive quantity-type is time(-difference) and there are two spaces, one representing a time interval of, for example, July and the other for example, the continent of Europe. On the first space there is a particular extensive quantity of this type known as duration. A particular journey might be a map (in an appropriate distributive category) from the first space to the second, hence via the functor the journey acts on the duration to produce on the continent a variable extensive quantity known as the sojourn (in each given part of the continent) of my journey. As another example, if I project my room onto the floor, the quantity of smoke is transformed into the quantity of smoke *over* the floor.

A further determination is suggested by the idea "space of quantity" which lies at the base of (not only cartesian coordinatizing but also) calculus of variations and functional analysis: the variable quantities (extensive or intensive) of a given type over a given space should themselves form a space (often infinite-



dimensional) which contains its own processes of "becoming" (continuous, differentiable, etc.) and is itself the domain of further variable quantities. This idea can be realized as follows: over a given distributive category of spaces, consider the linear category of all spaces equipped with given additions and all maps which preserve these; the forgetting functor from the latter to the former expresses in a general way that these quantity-types "are" spaces. But then in particular an extensive quantity-type from the distributive category to *this* linear category can be subjected to the further requirement that it be enriched (or strong) in the sense of enriched category theory, i.e. roughly that as a functor it be concordant with "becoming" (parameterization).

By contrast an intensive quantity-type is a *contravariant* functor, taking co-products to products, from a distributive category, but now a functor whose values have a multiplicative structure as well as an additive structure. Frequently the values of an intensive type are construed to be rigs, such as the ring of continuous or smooth functions or the lattice of propositional functions on the various spaces in the distributive category, with the functoriality given by substitution; however, since we also need to consider vector- and tensor-valued "functions", it is more adequate to consider that a typical value of an intensive quantity-type is itself a linear category, with composition in the latter being the multiplicative structure and with each spatial map inducing via the type a linear functor (in the opposite direction) between the two "petit" categories of intensive quantities on the domain and codomain spaces of the map. From the latter point of view the rigs are just endomap objects of certain preferred objects in these intensive categories, and in some examples (such as the analytic, though not the differentiable, study of projective space), knowledge of the rigs may not suffice to determine the intensive categories.

To exemplify the contravariant functoriality, the terminal map from a given space induces the "inclusion" of constant quantities of the given type as special "variable" intensive quantities on the space, while a given point of the space induces the *evaluation* at that point of any intensive quantity (caution: in general an intensive quantity may not be determined by the ensemble of its values at points); a particular journey of a month through a continent induces a transformation of any intensive quantity on the continent (such as the frequency with which a given language can be heard) into an intensive quantity varying over the month.

Again by specializing to the linear objects *in* the given distributive category as possible map-objects for the intensive categories assigned to each space, the important "space of quantity" idea, as well as a further enrichment requirement on the types, can also be realized for *intensive* quantities.

Moreover, if the distributive category is actually "cartesian-closed" (so has a "space of maps" between any two spaces, satisfying the objective relations which were used since the first days of the calculus of variations and which in this century were subjectively codified as "lambda-calculus") then the further important idea of the possible *representability* of components of an intensive quantity-type comes into play. Namely, the represented intensive quantity-type is defined to have as objects always the linear spaces in the distributive category itself, but each given space is defined to have as the map-objects of the corresponding intensive category the space of all maps from the given space to the spaces of linear maps between given linear spaces, the latter being the "representors"; an intensive quantity type is called representable if it is equivalent to a full part of this represented one. For example, the usual ring of smooth functions is representable when the constant scalars form a smooth space, and the lattice of propositional functions is representable when truth-values form a space (as they do in a topos).

It should be pointed out that there is a second doctrine of extensive/intensive quantities which agrees with the above when only "compact" spaces are considered, but which in general permits only "proper" spatial maps to induce (co-and contra-variantly) maps of quantities. Since they admit "totals", the extensive quantities which I described above should perhaps be thought of as being restricted to have "compact support", while the intensive quantities are "unrestricted" and thus might be representable, both of these features being compatible with my requirement of functoriality on *arbitrary* spatial maps in the distributive category. By contrast, the second "proper" doctrine is useful when considering "unrestricted" *extensive* quantities (such as area on the whole plane) but must correspondingly impose "compact support" restrictions on the *intensive* quantities, making the latter *non-representable*. These remarks presuppose the *relation* between extensive and intensive quantities, to which I will now turn.

The common spatial base of extensive and intensive quantities also supports the relation between the two, which is that the intensives act on the extensives. For example, a particular density function acts on a particular volume distribution to produce a resulting mass distribution. Thus it should be possible to "multiply" a given extensive quantity on a certain space by an intensive quantity (of appropriate type) on the same space to produce another extensive quantity on the same space. The definite *integral* of the intensive quantity "with respect to" the first extensive quantity is defined to be the *total* of this second resulting extensive quantity. This action (or "multiplication") of the contravariant on the covariant satisfies bilinearity and also satisfies, with respect to the multiplicative structure within the intensive quantities and along

any inducing spatial map, an extremely important strong homogeneity condition which so far has carried different names in different fields: in algebraic topology this homogeneity is called the "projection formula", in group representation theory it lies at the base of "Frobenius reciprocity", in quantum mechanics it is called "covariance" or the "canonical commutation relation", while in subjective logic it is often submerged into a side condition on variables for the validity of the rule of inference for existential quantification when applied to a conjunction.

It is in terms of such "action" ( or "multiplication") of intensive quantities on extensive quantities that the role of the former as "ratios" of the latter must be understood. As in the study of rational functions and in the definition of derivative, algebra recognizes that multiplication is fundamental whereas "ratio" is an inverse process; while the simple prescription "you can't divide by zero" may suffice for constant quantities, its ramifications for variable quantities are fraught with particularity, as reflected in even the purely algebraic "localization" constructions. For example, a given mass or charge distribution may not admit a density, with respect to volume, and not only the existence but also the uniqueness of such ratios may require serious study in particular situations, even though the multiplication which they invert is "everywhere" well-defined; the famous Radon-Nikodym theorem gives conditions for this in a specific context.

How can systems of extensive and intensive quantities, with action of the latter on the former, be realized on various distributive categories which mathematically arise? As mentioned above, the intensive quantities are often representable (indeed more often than commonly noticed, for example differential forms can be represented via the "fractional exponentiation" which exists in certain gros toposes). An important class of extensive quantities can be identified with the (smooth linear) *functionals* (with codomain a fixed linear space such as that of constant scalars) on the given intensive quantities, i.e. a distribution may sometimes be determined by the ensemble of all definite integrals (with respect to it) of all appropriate intensive quantities. This identification, supported in a particular context by the classical Riesz representation theorem (and in the homotopical context of section 2 below, by the universal coefficient theorem), contributed to the flourishing of functional analysis, but perhaps also distracted attention from the fact that extensive quantities are at least as basic as the intensive ones. At any rate, the fundamental projection formula/canonical commutation relation is automatic for those extensive quantities which can be identified as functionals on the intensive ones; here the action is *defined* in terms of the integral of the multiplication of intensive quantities.

This automatic validity of the fundamental formula holds also for a certain "opposite" situation in which a concept of intensive quantity can be defined to consist of transformations on given extensive concepts. More precisely, recall that I suggested above a general definition of extensive quantity type on a given distributive category as an enriched additive covariant functor from the given distributive category to the linear spaces in it. Given two such functors, we can consider *natural transformations* from one to the other, which thus can tautologously "multiply" extensive quantities of the first type to yield extensive quantities of the second type. Such natural transformations, however, are *constant* intensive quantities (i.e. "varying" only over the terminal space) since they operate over the whole distributive category. But the idea of natural transformation also includes all *variable* intensive quantities over some given space (and between two extensive functors), if we only make the following modification. An extremely useful construction, first emphasized by Grothendieck around 1960 (although it occurs already in Eilenberg and Mac Lane's original paper), associates a new category to any given object in a given category by considering as new objects all the *maps* with codomain the given object, and as new maps all the commutative triangles between these; this construction, a special case of the ill-named "comma category", has manifold applications revolving around the idea that both a *part* (with "multiplicity") of the given space as well as a *family* of spaces ("the fibers") smoothly parameterized by the given space are themselves objects in a new category; borrowing from Grothendieck, we may for short call this category the "gros" category of the given space (the "gros" category of the *terminal* space reducing to the given distributive category). Often a distributive category is in fact *locally* distributive, in the sense that for each space in it the associated "gros" category is again distributive. (The "petit" category of a space is usually a certain full subcategory of its "gros" category). A map between two spaces obviously induces by composition a coproduct preserving functor from the "gros" category of the first to the "gros" category of the second; in particular, the "gros" category of a space thus has a forgetting functor to the original distributive category of spaces. Composing this forgetting functor with two given extensive types, an intensive quantity varying over the given space may then be defined to be any natural transformation between the resulting composite functors. Thus according to this point of view, in the intensive category associated to a space, not only are the *maps* identified with intensive quantities varying over the space, but the *objects* are (or arise from) the types of extensive quantity which the whole category of spaces supports.

The most fundamental measure of a thing is the thing itself. If we replace "thing" by "object" (for example object in a category of spaces), then "itself"

may be usefully identified with the Pythagoras-Steiner-Burnside abstraction process discussed earlier: that is, isomorphic objects are identified, but all other maps are temporarily neglected. This obviously depends on what category the object is in, and the maps still play an important role in constructing and comparing new categories upon which the same abstraction process can be performed, notably the "gros" or "comma" categories of given spaces (as discussed above) and various "petit", "proper", "covering", "subobject" etc. subcategories of these. Moreover, in any locally distributive category there is for each map a "pullback" functor between the associated pair of "gros" categories, right adjoint to the obvious composition/forgetting functor previously mentioned. Thus, given a class of objects closed under coproduct (for example the class of finite, or discrete, or compact objects, or of the objects of fixed dimension, or intersections of these classes, etc) one can define a corresponding extensive quantity-type by assigning to each space (the abstraction of) the part of the "gros" category of that space which consists of those maps whose domains are in the class; this is obviously covariantly functorial via the composition/forgetting procedure. Given two such classes of objects, an intensive quantity from the one to the other, varying over a given space, can be defined to be (the abstraction of) any object of the "gros" category of the space which is *proper* in the sense that pulling back by it takes extensives of the one class into extensives of the other. Both the contravariant functoriality of these intensives as well as (tautologously) their action on such extensives is given by pullback, and the projection formula/CCR results from simple general lemmas about composition and pullback valid in any category. This *concrete* doctrine of quantity is explicitly or implicitly used in many branches of geometry, and I suspect that its direct use in many applications would be easier than translating everything into numbers (I recall a restaurant in New York in which customers, cooks, waiters, and the cashier may speak different languages, yet rapid operation is achieved without any written orders nor bills by simply stacking used dishes according to shape). One of the unsolved problems of the foundations of mathematics is to explain how and where the usual smooth distributions and functions of analysis can be obtained in this concrete mode.

As already the Grassmann brothers understood, the basic subject-matter of narrow-sense logic is quantities which are additively idempotent. The *intensive* aspect of this has been much studied, and is (at least fundamentally) concrete in the above sense, corresponding to parts without multiplicity (i.e. to subobjects); indeed one of the two basic axioms of topos theory is that subobjects are representable by (indicator maps to) the truth-value space. On the other hand the great variety of useful *extensive* logic has been little studied (at least

as logic). In practice logic is not really a starting-point but rather the study of *supports and roots* of non-idempotent quantity: for example, the inhabited *part* of the world is the part where population exists, yet population (unlike the indicator of the part) is a non-idempotent quantity; distributions have supports and a pair of functions determines the ("root-") space of their agreement as well as the ("open") subspace of their disagreement. While the Dedekind definition of a real intensive quantity as an ensemble of answers to yes-no questions has many uses, we should not let pragmatism blind us to the fact that a procedure for coming to know the quantity is by no means identical with the objectively operating quantity itself. The (still to be studied) extensive logic should be the codomain of an adequate general theory of the supports of extensive quantities, a theory accounting for certain rules of inference as reflections of the commutation relations for variable quantities; such relationships are studied in the branch of algebraic geometry known as intersection theory, but raising certain aspects of the latter to the level of philosophy should help to make them more approachable and also to suggest in what way they might be applied to other distributive categories.

It may be that, to accord more accurately with historical philosophical terminology, all the above occurrences of the word "quantity" should be replaced by "number", with the former being reserved for use in conjunction with the "affine" categories whose study has recently been revived by Schanuel, Carboni, Faro, Thiébaud and others; Grassmann seems to insist that numbers are *differences* of quantities (as for example work is a difference of energies, and duration a difference of instants), and further understanding of affine categories may reveal them as an objective basis of the link between distributive and linear categories. There are moreover "non-commutative affine objects" known as "symmetric spaces" which include not only Lie groups, but also spheres, but whose intrinsic categorical property and role has been little explored.

## 2. Homotopy Negates yet Retains Spatiality

The role of space as arena of "becoming" has as one consequence a quite specific form of the transformation of quantitative into qualitative; the seemingly endless elaboration of varied cohomology theories is not merely some expression of mathematicians' fanatical fascination for fashion, but flows from the necessity of that transformation.

One of the main features which distinguish the general "gros" spatial categories from the particular "petit" ones is the presence of spaces which can act as internal parameterizers of "becoming". Formally, the essential properties of

such a parameterizer space are that it is connected and strictly bipointed. Connectedness means that the space is not the coproduct of smaller spaces; when the category has a subcategory of "discrete" spaces, the inclusion functor having a left adjoint "components" functor, then connectedness of a space means that its space of components is terminal. Strict bipointedness means that (at least) two points (maps from the terminal space) of the space can be distinguished, where "distinguished" is taken in the strong sense that the two points are disjoint as subobjects, or in other words their "equalizer" is initial. (These definitions are often usefully extended from objects to "cylinders", i.e. to maps (with not-necessarily-terminal codomains) with a pair of common "sections" (generalizing "points")).

In order to maintain the rather heroic avoidance, which this paper has so far managed, of the traditional use of symbols to multiply the availability of pronouns, I will refer to the points of such a strictly bipointed connected space as "instants", without implying that that space is "one-dimensional" nor any further analysis of time. Note that, depending on the nature of the ambient "gros" distributive category, connectedness need not imply that the object is infinite; for the planning of activities and of calculations in the continuous material world, finite combinatorial models of the latter are necessary, and such models may in themselves constitute an appropriate category, the category of "simplicial sets" being a widely-used example.

Now a specific process of "becoming" in a certain space will be (accompanied by) a specific map to that space from a connected strictly bipointed space, whereby in particular the point to which one distinguished instant is mapped "becomes" the point to which the other distinguished instant is mapped. In particular, one map between two given spaces can "become" a second such map, as is explained by the usual definition of "homotopy" between the two maps, or equivalently (in the Hurewicz spirit), using if necessary the technique of embedding in presheaf categories to construe any distributive category as (part of) a cartesian-closed one, by applying directly the above account of "becoming" to points in the appropriate map-space. Obviously a very important application (of such internalization to a spatial category of the notion of "becoming") is the detailed study of dynamical processes themselves, bringing to bear the rich mathematical content which the category may have. However, in this section I will concentrate on the qualitative structure which remains after all such connected processes of "becoming" are imagined completed, that is, after any two maps which can possibly become one another are regarded as identical.

The traditional description of quantity as that which can be increased and decreased leads one to define a space as "quantitative" if it admits an action of a

connected strictly bipointed object, wherein one of the two distinguished instances acts as the identity on the space, whereas the other acts as a constant; thus the whole space can be "decreased to a point". This is a much stronger requirement than connectedness of the space and is usually called *contractibility*. This use of "quantitative" is not unrelated to the use of "quantity" in section 1, since representing objects for intensive quantities are often contractible.

With a "gros" spatial category it is usually possible to associate a new category called its *homotopy category*, in which homotopic maps become equal and contractible spaces all become isomorphic to the terminal space; in general two spaces which become isomorphic in the homotopy category are said to have the same homotopy type. For this association to exist the composition of homotopy-classes of maps must be well-defined, which in turn rests on an appropriate compatibility between connectedness and categorical product; in particular the product of two connected spaces should be connected. The latter is almost never true for "petit" distributive categories. In case the category is cartesian closed and has a components functor, the appropriate compatibility is assured if the components functor preserves products. To any such case Hurewicz's definition of the homotopy category, as the one whose map-spaces are the component spaces of the original map-spaces, can immediately be generalized, and indeed also extended to any category enriched in the given spatial category, such as pointed spaces, spaces with given dynamical actions, etc. yielding corresponding new qualitative categories which are enriched in the homotopy category. Using the product-preservation of the components functor and the fact that composition in a cartesian-closed category can be internally construed as itself a map whose domain is a product (of two map-spaces), it is easy to see that Hurewicz's definition supports a unique reasonable definition of composition of the maps in the homotopy category.

Now the main point which I wish to make is that essentially the whole account of space vs. quantity and extensive vs. intensive quantity given in section 1 reproduces itself at the qualitative level of the homotopy category. The latter is itself again a distributive category, cartesian closed if the original spatial category was. Indeed more precisely, the homotopy-type functor connecting the two actually preserves products, coproducts, and the map-space construction. On the other hand the homotopy category is *not* locally distributive: the passage to the parametric homotopy category of the "gros" category of a given parameter space seems to involve a further qualitative leap, not as passive as the corresponding passage in the quantitative context.

Although obtained by nullifying "quantitative" spaces, the homotopy category still admits extensive measurements of its objects, the most basic ones being the number of holes of given dimensions. The extensive quantity-types



here are usually called homology theories. Dually, the intensive quantity-types are cohomology theories, enjoying the features of contravariance, multiplicativity ("cup" product), and action as "ratios" of homology quantities. By a celebrated theorem, cohomology is often *representable* on the homotopy category, by objects known for their discoverers as Eilenberg-Mac Lane spaces. There is a strong tendency for basic homology and cohomology quantities to be (approximated by) linear functionals on each other. A new feature (probably distinguishing homotopy categories from the other distributive categories which do contain "becoming"-parameterizers and "quantitative" objects, although an axiomatic definition is unclear to me) is the appearance of *homotopy groups*, extensive quantity-types finer than homology and (co-) *representable* (by spheres). Note that the definition of "point" when applied in a homotopy category in fact means "component".

### 3. *"Unity and Identity of Opposites" as a Specific Mathematical Structure; Philosophical Dimension*

Not only should considerations of the above sort provide a useful guide to the learning and application of mathematics, but also the investigation of a given spatial category can be partly guided by philosophical principles. One of these is described, in conjunction with a particular application, in my paper "Display of graphics and their applications, as exemplified by 2-categories and the Hegelian 'Taco'".

Namely, within the system of subcategories of the category to be investigated, one can find a structure of ascending richness which closely parallels that of Hegel's *Science of Logic*, with each object to be investigated having its reflection and coreflection into each of the ascending levels. Here a level is formally defined as a functor from the given category (to a "smaller" one) which has both left and right adjoint sections; these sections are then the full inclusions of two subcategories which *in themselves* are "identical" (to the smaller category) but which as *subcategories* are "opposite" in the perfectly precise sense given by the adjointness, and the two composite idempotent functors resulting on the given category provide (via the adjunction maps) the particular reflection and coreflection in this level of any given space. In combinatorial topology such a level is exemplified by all spaces of dimension less than  $n$ , with the idempotent functors being called  $n$ -skeleton and  $n$ -coskeleton; in other cases the "dimensions" naming the levels may have a structure more (or less) rich than that of just natural numbers. Dimension "minus infinity" has the initial object and the terminal object as its two inclusion functors (in themselves, both are identical with the terminal category, but in the spatial

category they are opposite); it seems to me that there is good reason to identify the initial and terminal objects with Hegel's "non-being" and (pure) "being" respectively. While the relation of one level's being lower than another (and hence of one "dimension's" being less than or equal to another) can be defined in an obvious categorical way, the special nature of the levels subjects them to the further relation of being "qualitatively lower": namely, one level is qualitatively lower than another level if both its left and right adjoint inclusions are subsumed under the single *right* adjoint inclusion of the higher level. In many examples there is an *Aufhebung* or "jump": for a given level there is a smallest level qualitatively higher than it.

The *Aufhebung* of dimension "minus infinity" is in many cases "dimension 0", the left adjoint inclusion providing the discrete spaces of "non-becoming", the opposite codiscrete spaces forming an identical category-in-itself which is however now included as the chaotic "pure becoming" in which any point can become any other point in a unique trivial way. Both initial and terminal objects are codiscrete. This level zero (in itself) is often very similar to the category of abstract sets, although (for example in Galois theory) it may not be exactly the same; as I tried to explain in my 1989 Cambridge lectures, the double nature of its inclusion into mathematics may help to resolve problems of distinguishability vs. indistinguishability which have plagued interpretation of Cantor's description of the abstraction process, (and hence obscured his definition of cardinals). This discrete/codiscrete level is often special in the further respect that its left (discrete) inclusion functor has a further left adjoint, the "components" functor for the whole category (which, as discussed above, should further preserve finite products).

The *Aufhebung* of dimension zero strongly deserves to be called dimension one: its equivalent characterization, as the smallest level such that any space has the same components as the skeleton (at that level) of the space, has the clear philosophical meaning that if a point (or figure) can become another one, then it can do so along some 1-dimensional process of "becoming". Here dimensionality *of a space* (such as a parameterizer) is defined negatively in terms of skeleta (rather than "positively" in terms of coskeleta which are typically infinite dimensional).

For the levels qualitatively higher than zero, the *right* adjoint inclusion also preserves *co*-products, a very special situation even for topos theory. In a topos having a "becoming"-parameterizer, the truth-value object itself is contractible (as pointed out by Grothendieck), permitting "true" to become "false" in a way overlooked by classical logic; hence the 1-skeleton of the truth-value space presents itself as a canonical (though perhaps not adequate) parameterizer of "becoming" or "interval". These suggest only a few of the many open prob-

lems, involving calculation of the many examples, which need to be elaborated in order to clarify the usefulness of these particular concrete interpretations of the dialectical method of investigation. We very much need the assistance of interested philosophers and mathematicians.

### References

- [1] Carboni, Aurelio, "Affine Categories", *Journal of Pure and Applied Algebra*, 61 (1989) 243–250.
- [2] Dress, Andreas. In: *Algebraic K-Theory II*, Springer Lecture Notes Mathematics, 342, 1973.
- [3] Faro, Emilio, "Naturally Mal'cev Categories", to appear in *Proceedings of 1990 Como International Conference on Category Theory*.
- [4] Grassmann, Hermann, *Ausdehnungslehre*, 1844.
- [5] Grothendieck, Alexander, *Pursuing Stacks* (manuscript), 1983.
- [6] Kelly, Gregory Maxwell, *Basic Concepts of Enriched Category Theory*, Cambridge UP, 1982.
- [7] Kelly, G. M. / Lawvere, F. W., "On the Complete Lattice of Essential Localizations", *Bulletin de la Société Mathématique de Belgique*, XLI (1989) 289–319.
- [8] Lawvere, F. William, "Elementary Theory of the Category of Sets", *Proceedings of the National Academy of Sciences, USA*, 52 (1964) 1506–1511.
- [9] —, "Adjointness in Foundations", *Dialectica* 23 (1969) 281–296.
- [10] —, "Metric Spaces, Generalized Logic and Closed Categories", *Rendiconti del Seminario Matematico e Fisico di Milano* 43 (1973) 135–166.
- [11] —, "Toward the Description in a Smooth Topos of the Dynamically Possible Motions and Deformations of a Continuous Body", *Cahiers de Topologie et Géométrie Différentielle*, XXI-4 (1980) 377–392.
- [12] —, "Introduction" to *Categories in Continuum Physics*, Springer Lecture Notes in Mathematics 1174, 1986.
- [13] —, "Taking Categories Seriously", *Revista Colombiana de Matemáticas* XX (1986) 147–178.
- [14] —, "Categories of Spaces May Not be Generalized Spaces, as Exemplified by Directed Graphs", *Revista Colombiana de Matemáticas* XX (1986) 179–185.
- [15] —, "Display of Graphics and their Applications, as Exemplified by 2-Categories and the Hegelian Taco", *Proc. of the First International Conference on Algebraic Methodology and Software Technology*, Univ. of Iowa, 1989.
- [16] Mac Lane, Saunders, "Duality for Groups", *Bulletin of the American Mathematical Society* 56 (1950) 485–516.
- [17] Schanuel, Stephen, "Dimension and Euler Characteristic", *Proceedings of 1990 Como International Category Theory Meeting*, Springer Lecture Notes in Mathematics 1488, 1991.
- [18] Thiébaud, Michel, "Modular Categories", *Proceedings of 1990 Como International Category Theory Meeting*, Springer Lecture Notes in Mathematics 1488, 1991.

# Structural Analogies Between Mathematical and Empirical Theories

ANDONI IBARRA (San Sebastian) / THOMAS MORMANN (Berlin)

## 1. Introduction

Time and again philosophy of science has drawn analogies between mathematics and the empirical sciences, in particular physics. The orientation of these analogies, however, has been rather different. In the heyday of Logical Positivism philosophy of mathematics was considered the model for philosophy of the empirical sciences.<sup>1</sup>

For some time one can witness the opposite approach: the import of ideas from the realm of philosophy of empirical science to the realm of philosophy of mathematics. We would like to point out the following approaches of such a transfer:

### *Methodological Analogy*

Lakatos proposed to transfer his "Methodology of scientific research programs" ("MSRP") from the sphere of empirical science to the realm of mathematics. He claimed that there exists a *methodological* parallelism between the empirical and ("quasiempirical") mathematical theories: in both realms one uses the method of "daring speculations and dramatic refutations".<sup>2</sup>

### *Functional Analogy*

Quine proposed a functional analogy between mathematical and empirical knowledge. His approach is based on the holistic thesis that mathematical concepts like "numbers", "functions" or "groups" play in the global context of

---

<sup>1</sup> This leads to deductively oriented conceptions of empirical science, in particular to the so called "received view", cf. (Nagel 1961), (Suppe 1974).

<sup>2</sup> Cf. (Lakatos 1978, p. 41). There are not too many authors who have continued Lakatos' approach despite its very high esteem in all quarters of philosophy (and history) of mathematics. We mention (Howson 1979), (Hallett 1979), and (Yuxin 1990).

scientific knowledge essentially the same role as physical concepts like "electron", "potential" or "preservation of symmetry".<sup>3</sup>

We do not want to criticise these two approaches in detail here, we simply claim that neither Lakatos' "methodology of scientific research programmes" nor Quine's holism has to be the last word in this philosophical issue. We would like to sketch another analogy between empirical and mathematical knowledge, namely, a *structural analogy between empirical and mathematical theories*. It might not be incompatible with those mentioned above. In particular, it may be considered as a specification of Quine's holistic approach. The structural analogy between mathematical and empirical theories, which we want to explain, is based on the general thesis that cognition, be it empirical, mathematical or of any other kind, e.g. perceptual, has a *representational structure: Cognition is representation*.

This thesis can be traced back at least to Kant who maintained that cognition is governed by representational structures, e.g. by the forms of intuition (space and time), and certain categories of understanding like causality. Among its more recent adherents we may mention C. S. Peirce and E. Cassirer. But we do not want to deal here with the problem of general representational character of cognition from a transcendental philosophical viewpoint as Kant did. We would rather like to make plausible that cognition is representation by presenting an inductive argument.

First of all, let us recall that the representational character of cognition is not restricted to scientific knowledge but pervades all kinds of cognition, e.g. *perception* and *measurement*. For one reason or another, the similarity between these kinds of cognition and scientific theories is often underestimated or even denied: perception seems to belong to the merely subjective sphere of the individual, and measurement, though it may be objective, seems to lack a theoretical component. Hence, neither perception nor measurement seem to have much in common with scientific cognition. This impression is wrong. They all share a common representational character and this feature is of crucial importance for their epistemological structure.

Now, by ascribing a representational structure to perception, measurement, and scientific knowledge we do not want to claim that the representational structure in all of them is the same. This does not seem plausible. It may well be that the representations in these different areas are determined by quite different constraints, cf. (Goldman 1986, part II). This, however, does not exclude the possibility that they all should be dealt with a comprehensive epistemology which investigates cognition in all its different realizations and

---

3 Cf. (Quine 1970), also (Putnam 1979), and (Resnik 1988).

treats them from a unified perspective.<sup>4</sup> Let us mention some trends towards that epistemology.

### *Measurement as Representation*

Measurement is representation in a quite direct and intuitive sense, namely, *measurement is representation of empirical facts and relations by numerical entities and relations*. The explication of this elementary idea in the framework of the so called *representational theory of measurement* has been developed, among others, by Stevens and Suppes, cf. (Suppes 1989), (Mundy 1986). This has led to a comprehensive classification of the various kinds of measurement scales and their transformation groups and invariants. It can be considered as an empirical analogue to Klein's *Erlangen Programme* which aimed to classify the different geometries. This analogy often has been recognized but only recently have there been serious attempts to consider the representational theory of numerical measurement, Klein's classification of geometries, and the various attempts of a general "geometrization" of physics as special cases of a single coherent "general theory of meaningful representation", cf. (Mundy 1986).

### *Perception as Representation*

A perceived object is the result of a representational, constructive process. This is shown very clearly by the various phenomena of perceptual constancy, for example color or gestalt constancy. Up to now, there is no unanimity among the different approaches of cognitive psychology and cognitive science of how the representational character of perception is to be understood precisely. The so called "bottom-up" and the "top-down" approaches conceptualize it in a quite different way, cf. (Goldman 1986). Nevertheless, practically all sciences dealing with perception agree (or at least are compatible) with the assertion that some kind of representation and symbolic construction is involved. Following Gödel an alleged analogy of "mathematical perception" and "visual perception" often has been taken as an argument for a robust platonism or realism which claims that the mathematicians "perceive" mathematical objects just as ordinary people perceive the more mundane things of the ordinary world.

This analogy of mathematical and ordinary perception may subjectively be

---

<sup>4</sup> For general considerations for such a comprehensive epistemology, cf. (Goodman / Elgin 1988, p. 16). It should be noticed, however, that already more than sixty years ago Cassirer set about the project of embedding philosophy of science in a general representational theory of symbolization in his *Philosophie der symbolischen Formen*, cf. (Cassirer 1985).

quite justified, i.e. it may be the case that mathematicians in the course of their research believe that they "perceive" mathematical objects as they "really are" just as the layman believes that he perceives the table in front of him just as it "really is". But the cognitive sciences teach us that – *pace* Gibson – this is a somewhat simplistic account of perception. Hence, it would be interesting to investigate the problem whether the analogy of mathematical and visual perception could survive as an argument for mathematical realism (platonism) if it is based on an updated account of perception.

### *Cognition as Representation*

Quite generally, Cassirer claims that the search for representational *invariants* is not restricted to perception but common to all kinds of cognition, be it perception, measurement, thinking or whatever else. He considers the tendency towards "objectification" in perception only as the rudimentary form of a general tendency in conceptual, in particular mathematical, thought, where it is developed far beyond its primitive stage (cf. also (Cassirer 1944, p. 20)):

"A critical analysis of knowledge reveals that the "possibility of object" depends upon the formation of certain invariants in the flux of sense impressions, no matter whether these be invariants of perceptions or of geometrical thought, or of physical theory" (Cassirer 1944, p. 21).

In the case of empirical and mathematical cognition the thesis is that empirical theories are representations, and correspondingly that mathematical theories are representations.

Thus, in order to clarify the representational character of empirical and mathematical cognition and to make plausible the structural analogy between empirical and mathematical theories we are led to the elucidation of the concept of *theory* in both realms of knowledge. This is one of the central conceptual problems of philosophy of science: to provide an adequate explication of this term.<sup>5</sup>

The outline of the paper is as follows: in the next section we sketch the representational character of empirical theories.<sup>6</sup> Then we want to show through the example of *group theory* in what sense a "typical" mathematical theory can

---

5 A large part of the criticism against Logical Empiricism can be formulated as criticism against the inadequate theory concept of this approach: an empirical theory simply is not a partially interpreted calculus as the so called "received view" holds it, and similarly a "real life" mathematical theory like algebraic topology or commutative algebra cannot be explicated adequately in terms of a calculus of meaningless formal signs.

6 We base our approach largely on Henry Margenau's "Methodology of modern physics", (Margenau 1935).

be considered as a representation in quite an analogous way as empirical theories are representations. Finally, we will point out how *category theory* can be considered as a useful formal tool for the representational reconstruction of mathematical and empirical theories. We will close with some remarks on the relation of our representational approach with some more traditional currents in the philosophy of mathematics.

## 2. Empirical Theories as Representations: Data, Symbolic Constructs, and Swing

To explain the thesis of the representational structure of empirical theories we start by distinguishing two levels of physical conceptualization. We thereby follow the approach of the philosopher and scientist Henry Margenau who among others distinguished between the *level of data* and the *level of symbolic constructs*.<sup>7</sup> He considers the following paradigmatic example:

"... we observe a falling body, or many different falling bodies; we then take the typical body into mental custody and endow it with the abstract properties expressed in the law of gravitation. It is no longer the body we originally perceived, for we have added properties which are neither immediately evident nor empirically necessary. If it be doubted that these properties are in a sense arbitrary we need merely recall the fact that there is an alternate, equally or even more successful physical theory – that of general relativity – which ascribes to the typical bodies the power of influencing the metric of space, i.e. entirely different properties from those expressed in Newton's law of gravitation" (Margenau 1935, p. 57).

It should be evident that this two-level-structure is not restricted to mechanics, it pervades all parts of physics.

Even if the realm of *symbolic constructs* in physics is *not* determined in the same rigid way as the realm of *data*, it is not completely arbitrary. There are general requirements concerning *symbolic constructs*:

"Physical explanation would be a useless game if there were no severe restrictions governing the association of constructs with perceptible situations.

---

<sup>7</sup> We rely on Margenau because his account is intuitive and avoids any unnecessary technical fuss. However, Margenau is not the only one, and not the first, who makes such a distinction: some more or less implicate remarks on the representational character of empirical theories can be found in Duhem's account of "The aim and structure of physical theory"; see especially (Duhem 1954, Ch. 8). In a formally very sophisticated manner the distinction between *data* ("Intended Applications") and *symbolic constructs* ("Models") is elaborated in the so called structuralist approach of philosophy of science, cf. (Balzer / Moulines / Sneed 1987).



For a long time it had been supposed that all permissible constructs must be of the kind often described as mechanical models or their properties, but this view is now recognized as inadequate. ... While no restriction can be made as to their choice, their use is subject to very strong limitations. It is easy to find a set of constructs to go with a given set of data, but *we require that there be a permanent and extensive correspondence between constructs and data* " (Margenau 1935, p. 64).

Putting together the ingredients of *data*, *symbolic constructs*, and their *correspondence* we propose the following general format of an empirical theory. According to the representational approach an empirical theory is a *representation* of the following kind:<sup>8</sup>

$$f: \mathcal{D} \longrightarrow C$$

The realm  $\mathcal{D}$  of *data* is *represented* via a mapping  $f$  by the realm  $C$  of *symbolic constructs*. The requirement that there must be a *permanent and extensive correspondence between constructs and data* is expressed by the requirement that the representing mapping  $f$  from  $\mathcal{D}$  to  $C$  cannot be just any mapping but has to respect the structure of  $\mathcal{D}$  and  $C$ . Therefore some constraints have to be put upon it which may not always be satisfiable. Thus, the thesis that a theory has a representational structure implicitly makes the claim that the  $\mathcal{D}$  can be represented by  $C$  in an appropriate way.<sup>9</sup> How this is to be understood precisely depends on how we conceptualize the realms of *data* and *symbolic constructs*.<sup>10</sup>

In philosophy of science, the specific nature and relation of these two levels of empirical theories have been a topic of much discussion. A rather popular account took  $\mathcal{D}$  as the observable and  $C$  as the non-observable. But this has not been the only approach. Others have considered  $\mathcal{D}$  as the empirical, and  $C$  as the theoretical. It cannot be said that unanimity has been achieved how these levels of conceptualization have to be understood precisely. Probably *The One and Only Right Explication* does not exist. In any case, the

---

<sup>8</sup> This is in some respect an oversimplification: as it turns out, a (mathematical or empirical) theory can be reconstructed as a whole bunch of representations of the above mentioned kind. Thus, more precisely we should consider a representation  $f: \mathcal{D} \longrightarrow C$  as the smallest meaningful element of a theory.

<sup>9</sup> This claim corresponds to the "empirical claim of a theory" of the structuralist approach, cf. (Balzer / Moulines / Sneed 1987) and the "theoretical hypothesis" of the state space approach of van Fraassen and Giere, cf. (van Fraassen 1989) or (Giere 1988).

<sup>10</sup> In the case of mathematical theories  $f$  will be a *functor* between the *Category  $\mathcal{D}$  of Data* and the *Category  $C$  of Symbolic Constructs*.

emphasis should not be laid upon what the constituents  $\mathcal{D}$  and  $\mathcal{C}$  "are in themselves"; what is important is the functional aspect of representation. Thus, the representational approach puts the essence of a theory neither in the things the theory talks about nor in the concepts used by the theory but, so to speak, in the interstices, in the ontologically ambiguous space of the representational relation between *data* and *symbolic constructs*.

We do not want to offer any argument for our specific position in this issue, but simply characterize our stance by the following remarks:

- (i) The distinction between *data* and *symbolic constructs* is no absolute distinction, i.e. in one context entities can work as *data* and in another context as *symbolic constructs*. In particular, *data* must not be considered as the "immediately given" of some Logical positivists. Hence, Margenau made the proposal to replace the term "*data*" by the less misleading term "*habita*", i.e. that what one has at his disposal or takes as the starting point of a research undertaking, cf. (Margenau 1935). In a similar vein Goodman points out, that "the given" or even "the immediately given" does not exist. Epistemology should consider "the given" as "the taken", i.e. as something which has been taken as the relative starting point or relative basis, (Goodman 1978, p. 10).
- (ii) It is an important task for the philosophical reconstruction of empirical and mathematical theories to explicate in a precise manner the *structure* and the *functions* of the *correspondence between data and symbolic constructs*.<sup>11</sup>

What is the representation of *data* by *symbolic constructs* good for? This is a very deep problem whose surface we can only scratch in the context:

- (a) *Symbolic constructs* generate a "conceptual surplus" which can be used for determining and predicting previously unaccessible aspects of *data*. For example, partially known kinetic data are embedded into the framework of *symbolic constructs* like *forces*, *Hamiltonians*, or *Lagrangians* in order to obtain new information not available without them.

---

<sup>11</sup> In the framework of the structuralist approach this correspondence is explicated in the following way: a theory  $T$  has a (more or less determined) domain  $I$  of "Intended Application" – Margenau's *data* – and a domain  $K$  of conceptual structures – Margenau's *symbolic constructs*. Thus  $T$  is an ordered couple  $T = \langle K, I \rangle$ . The global claim of  $T$  is that  $I$  can be embedded in  $K$  in such a way that a connected class of *data* corresponds to a connected class of *symbolic constructs*. The most detailed account of this approach presently available is (Balzer / Moulines / Sneed 1987).

- (b) *Symbolic constructs* have an explanatory function and serve to embed the *data* into a coherent explanatory theoretical framework. That is, the correspondence between *data* and *symbolic constructs* is the basis of physical explanation. To use once again the just mentioned example: a kinetic system may be explained causally by referring to theoretical constructs like *forces*.

Hence, physical explanation can be described as a movement of the following kind: it starts in the range of *data*, swings over into the field of *symbolic construction*, and returns to *data* again:

$$\mathcal{D} \longrightarrow \mathcal{C} \longrightarrow \mathcal{D}$$

More generally, we can characterize the activity of the scientists, be it explanation, prediction, or conceptual exploration, as an oscillating movement between the area of *data* and the area of *symbolic construction*. Following Margenau we want to call it *swing* – more philosophically oriented minds may even call it a *hermeneutic circle*. Somewhat more explicitly, this last expression may be justified as follows: we start with a limited and partial "Vorverständnis" of the *data*. Then the *data* are embedded and represented in an interpretatory framework of *symbolic construction* which may be used to yield a fuller understanding of the *data*.

The purpose of the *swing* is manifold: it may be used to obtain new information about the *data*, or to provide an explanation for them, or even to excite new conceptual research concerning the *symbolic constructs*.<sup>12</sup>

For the purpose of this paper we want to emphasize the following features of *data*, *symbolic constructs*, and *swing* :

*Relativity*: whether an entity *e* belongs to the realm of *data* or to the realm of *symbolic constructs* is context-dependent: in one context *e* may be considered as a *datum*, in another context as a *symbolic construct*.

*Plurality*: the realm of *symbolic constructs* is not uniquely determined by the realm of *data*: there may be several rival (incompatible) *symbolic constructs* for one and the same *data*.

*Usefulness, Economy and Explicitness*: the *symbolic constructs* are construc-

<sup>12</sup> It might be interesting to note that for perception some authors have proposed an analogous cycle between the "objects" perceived (*data*) and the "schemata" (*constructs*) structuring perception, cf. (Neisser 1976, p. 20f). In general, in cognitive science there is more or less agreement on the assertion that perception uses a mixture of "bottom-up" (*data*  $\Rightarrow$  *constructs*) and "top-down" (*constructs*  $\Rightarrow$  *data*) processes to realize the perceived object or situation, cf. (Goldman 1986, p. 187).

ted with certain purposes, they have to be useful. This usually implies economy and explicitness. To invoke an example of N. Cartwright taken from the empirical sciences:

" A good theory aims to cover a wide variety of phenomena with as few principles as possible. ... It is a poor theory that requires a new Hamiltonian for each new physical circumstance. The great explanatory power of quantum mechanics comes from its ability to deploy a small number of well-understood Hamiltonians to cover a broad range of cases, and not from its ability to match each situation one-to-one with a new mathematical representation. That way of proceeding would be crazy" (Cartwright 1983, pp. 144/145).

*symbolic constructs* in mathematical theories have to have the same or at least similar virtues. The merits of Poincaré's theory of fundamental groups reside in the fact it is possible to cope with a wide range of seemingly disparate phenomena through the concept of fundamental groups. The fact that each manifold has a fundamental group is in itself only of limited interest. In order to be useful, one must be able to effectively calculate this construct. And, equally important, it must turn out that the fundamental group reflects important traits of the spaces.

Quite generally, in mathematics *symbolic constructs* represent or express *the possible* contexts of *data*. According to Peirce's *Pragmatic Maxim* the possible contexts of *data* may be identified with their *meaning*:

"Consider what effects, that might conceivably have practical bearings, we conceive the objects of our conception to have. Then our conception of these effects is the whole of our conception of the object" (Peirce 1931/35 [5.402]).

This can be spelt as follows: the "practical bearings" a mathematical object might conceivably have are its (functional) relations to other mathematical objects. For example, the meaning of a particular manifold reveals itself in its possible relations with other manifolds or, more generally, with other mathematical entities and, as we will sketch in the following section for manifolds, an important part of these possible relations can be described in the framework of group theory. On the other hand, *category theory* can be considered as the realization of a kind of functional *Pragmatic Maxim* according to which the meaning of a mathematical object is to be seen in its relations to other mathematical objects. *category theory* does not care much about the objects of mathematical theories so that it is formally possible to eliminate them in favor of relations. Thus, from a category-theoretical point of view, an entity gets its meaning *not* by its underlying substance, i.e. its underlying sets of members or its internal properties, but through its external relations to other

objects of the category. Impressive examples of this fact are provided by the so called "arrow style" definitions. A simple case is the definition of the kernel of a group homomorphism. A more spectacular one is Lawvere's category theoretical reconstruction of the concepts of set and element through the concept of a subobject classifier which is a pure "arrow-style" concept and makes no use of any set theoretical concept, cf. (Goldblatt 1979).

### 3. *The Case of Mathematics: Data, Symbolic Constructs, and Swing in Group Theory*

The structural analogy between empirical and mathematical knowledge which we want to exhibit consists in the fact that the structure of mathematical theories and research can be explained in terms of *data*, *symbolic constructs*, and *swing* as well. Our example is a tiny part of group theory.

#### 3.1. *Groups as Symbolic Constructs*

We would like to explain the idea of groups as *symbolic constructs* through the example of the *fundamental group of manifolds* introduced by Poincaré at the turn of the century.

The *fundamental group* of a manifold is the prototype of a *symbolic construct* which plays a central role for the solution of the following general topological problem:

Given two manifolds  $M_1$  and  $M_2$  the problem is to prove that they are unrelated in the sense that there does not exist a continuous map  $f: M_1 \longrightarrow M_2$ . A famous case in this context is *Brouwer's fixpoint theorem* that can be considered as a paradigmatic example.

In order to solve such a problem it often turns out to be convenient, even necessary, to replace the manifolds themselves by appropriate *symbolic constructs*, in our simple case by their fundamental groups  $\pi_1(M_1)$  and  $\pi_1(M_2)$ , and to convert the geometrical problem into an algebraic one. That is to say, one shows that there does not exist an algebraic map, i.e. a group homomorphism, between the fundamental groups  $\pi_1(M_1)$  and  $\pi_1(M_2)$ . Then, due to the correspondence between manifolds and their fundamental groups, one can conclude that there does not exist a continuous map between the manifolds themselves.<sup>13</sup>

---

<sup>13</sup> Compared with the *Data* manifold the definition of the *Symbolic Construct* "fundamental group" is somewhat complicated. One may know the *Data* quite well, i.e. one may be able to identify or to distinguish them quite easily

Thus, in the same way as for empirical theories, in the case of Poincaré's theory of the fundamental group we have the two constituents: the *level of data* – the manifolds – and the *level of symbolic constructs* – the fundamental groups of the manifolds. Further, we have the *swing* from the manifolds and their geometric relations to the groups and their algebraic relations back to the manifolds. To put it bluntly: the proof of *Brouwer's fixpoint theorem* is essentially nothing but the *swing*.

Let us finally consider briefly the topic of plurality. In the later development of topology it turned out that the fundamental groups are by no means the only *symbolic constructs* for manifolds. A huge generalization is provided by the so called higher homotopy groups  $\pi_i(M)$  ( $i \geq 2$ ) for manifolds. The fundamental group is only the first of an infinite series of grouplike *symbolic constructs*.<sup>14</sup>

### 3.2. Symbolic Constructs of Groups

Usually new mathematical entities first appear as *symbolic constructs*.<sup>15</sup> However, once such a construct has been established in the mathematical discourse, it rather quickly takes the role of a datum for which further *symbolic constructs* are built. In the case of groups we want to consider the *symbolic construct* of characters related to a group  $G$ .

A character of a group  $G$  is a representation of  $G$  into the complex numbers  $\mathbb{C}$ , i.e. a function from  $G$  into  $\mathbb{C}$  with certain special properties we need not consider in any detail. The *symbolic construct* of the set of characters  $\mathbb{C}(G)$  forms a vectorspace and serves as a model of the group  $G$ , and can be used as a powerful tool to investigate its structure. A well known example is the *theorem of Burnside* which deals with the solvability of certain finite groups. For quite a long time the only available proof of this theorem made crucial use of the *symbolic construct* of characters, although the statement of *Burnside's theorem* can be formulated quite independently from this concept.

Again we are confronted with the typical *swing*: Starting from the level of *data*, in our case they are the set (or category) of groups, we move into the

---

without being able to calculate their fundamental groups. In the case of *Brouwer's Fixpoint Theorem* it is easy to calculate the fundamental groups of the manifolds involved.

<sup>14</sup> It is remarkable that these higher homotopy groups are not known completely even for quite "elementary" manifolds like the 2-dimensional sphere.

<sup>15</sup> Groups as *Symbolic Constructs* are introduced for the first time by Lagrange and Euler in the course of their investigations on quadratic forms and potential rests, cf. (Dieudonné 1976), (Wussing 1969).

field of *symbolic construction*, the set (or category) of vectorspaces, and return to the realm of groups again.

Here too we can witness a pluralism of *symbolic constructs*. Problems of groups in general, and problems of solvability in particular, *need not* be treated by characters. Much later, a proof of *Burnside's theorem* was found which does not depend on the *symbolic construct* of characters.

#### 4. Formal Tools: Category Theory

The representational reconstruction of mathematical theories sketched here for a small part of group theory has the advantage that a lot of work for its formal elaboration has already been done. To nobody's surprise it turns out that one can use the tools of *category theory* for the representational reconstruction of mathematical theories.

As is well known, the *fundamental group of Poincaré* can be considered as a functor  $\mathcal{P}$  from the *category of manifolds* to the *category of groups*:

$$\mathcal{P}: \mathcal{M} \longrightarrow \mathcal{G}$$

In a similar vein one can interpret the correlation of groups and characters as a functor  $\mathcal{C}$  from the *category of groups*  $\mathcal{G}$  to the *category of vectorspaces*  $\mathcal{V}$ :

$$\mathcal{C}: \mathcal{G} \longrightarrow \mathcal{V}$$

The fact that  $\mathcal{P}$  and  $\mathcal{C}$  are functors can be considered as a precise formulation of the requirement that there exists a permanent and extensive correspondence between *data* and *constructs*, i.e. relations between *data* have to correspond (at least partially) to relations between *constructs* and vice versa.<sup>16</sup>

According to category theory, *group theory* is a whole bunch of grouplike representations, or better, a net of grouplike representations. This net is to be conceptualized as an open net, i.e. as a net which is extended in different ways and directions: new knots are added, and new connections between already existing knots are constructed, and so on...

#### 5. Conclusion.

The representational approach might prove to be especially useful in coping with some weaknesses which the philosophy of mathematics traditionally suf-

---

<sup>16</sup> What this mean exactly depends on the specifics of the case in question, but it can be spelt out in the framework of a *general theory of meaningful representation*, cf. (Mundy 1986).

fered from, i.e. the exaggerated inclination to stick to attitudes like *elementarism*, *fundamentalism*, and *ontologism*.

*Elementarism* claims that it is sufficient to understand elementary mathematical theories as the arithmetic of natural numbers, thus gaining complete philosophical insight into the whole enterprise of mathematics. The concepts of category theory used in the representational approach are *technical* concepts which are rather immune to an utterly elementarist approach. Thus the representational approach is closer to the conception of the working scientist.

*Fundamentalism* maintains that the most important task of the philosophy of mathematics is to provide an absolutely secure foundation of mathematics. *Fundamentalism* localizes the philosophical problematic of mathematics in its foundations, be it logic, set theory, or any other foundational discipline. This leads to a strictly *hierarchical* and *global* organization of mathematical knowledge. Contrary to this (inadequate) conception of mathematics the representational approach favors a local, more flexible organization of mathematics as a *non-hierarchical* net of interrelated units.

Finally, *ontologism* concentrates rather exclusively on global questions like the following: What "mode of being" pertains to mathematical entities? To this question there exists a whole spectrum of answers. On the one end, we find a solid platonism which assigns mathematical objects to an exclusive area whereas, on the other end, we find eliminative conceptions which try to reinterpret the domain of mathematics nominalistically or physicalistically. They hope to get rid of the ontological problems of mathematics once and for all. Between those extreme positions we find constructivist approaches which put various constraints on mathematical entities. In principle we do not think these approaches to be wrong but we would like to remark that these global ontological claims of philosophy towards mathematics appear, from a naturalistic perspective, to be rather strong. They do not correspond to anything in philosophy of empirical sciences. The question "What is an electron?" sounds strange while philosophers of mathematics frequently ask "What is a number?" and similar questions.<sup>17</sup>

The *relative, context-dependent* characterization of mathematical entities as *data* and *symbolic constructs*, however, leads the representational approach to a *distributed* and *variegated* ontology of the objects of mathematical discourse. One cannot assume that the mode of being of entities belonging to different levels of conceptualization is the same and remains fixed once and for all.

---

<sup>17</sup> Often, this approach is related to *fundamentalism* in localizing the ontological question in a fundamental domain, e.g. in arithmetics of natural numbers or set theory.



Rather, the ontological status of mathematical entities is a *variable* and depends on the development of mathematics.

As Margenau pointed out already 50 years ago a simplistic *Yes/No* attitude concerning the existence of scientific entities is inappropriate:

"Do masses, electrons, atoms, magnetic field strengths, etc., *exist*? Nothing is more surprising indeed than the fact that ... most of us still expect an answer to this question in terms of yes and no. ... Almost every term that has come under scientific scrutiny has lost its initially absolute significance and acquired a range of meaning of which even the boundaries are often variable. Apparently the word *to be* has escaped this process" (Margenau 1935, p. 164).

Even if we assume for the sake of the argument that the isolated claim " $\pi$  exists" makes sense, one must be a very hardheaded platonist to maintain that  $\pi$  exists *in the same way* as, say, an entity like an "extraordinary cohomology theory" – a rather complex entity which nevertheless can be made the object of study.

It should be noted that the problem of ontological diversification is not a special problem of the empirical sciences. It concerns the social sciences and common sense knowledge as well: Does it really make sense to maintain that objects like "San Sebastian", "the European Community", or "the development of capitalism in the 20th century" exist in the same way as the notorious apple tree in the philosopher's garden?

Thus, taking into account the structural analogy between empirical and mathematical theories as representations, we maintain that mathematics *shares* with empirical science this feature of a variegated ontological status of its objects. That is, we maintain that Margenau's remarks concerning the *blurred ontology* status of scientific entities also apply to mathematics. Regrettably this *common ground* of mathematics and empirical science is rarely recognized by philosophers of mathematics. In the realm of physics, for example, philosophers of mathematics often take a robust realism for granted, thereby accepting an artificial wall between mathematics and empirical science. The representational structure of empirical *and* mathematical theories, however, renders it dubious that there is a sharp and clear ontological distinction between the physical and the mathematical. At this point the representational approach meets Quine's holistic account of science, cf. (Resnik 1988).

Hence, taking into consideration this common feature of mathematics and empirical science it is evident that the philosophy of mathematics cannot restrict itself to the "foundations" or the "elements" of mathematics. It has to

pay attention to the ongoing process of mathematical progress.<sup>18</sup>

Stressing the *common ground* of empirical and mathematical cognition as it is exhibited in the representational structure of mathematical and empirical theories, the philosophy of mathematics could tap the sources of present day philosophy of empirical science. There, problems of ontology are dealt with in a far more liberal and sophisticated way than in the Logical Empiricism of the thirties. Nowadays, the ontological diversification of scientific entities is widely recognized, as is witnessed by the talk of causality, possibilities, or counterfactuals, even in quarters which consider themselves to belong to the analytic tradition. In particular, one may consider the ongoing debate on realism as an effort to overcome the far too simple dualism of "does exist" versus "does not exist".<sup>19</sup>

Up to now, however, philosophy of mathematics seems to have ignored this debate. In order to gain contact again with the rest of philosophy of science, philosophy of mathematics has to give up the pernicious concentration on such idiosyncratic "-isms" as *elementarism*, *fundamentalism* and *ontology*.

We hope the representational approach can be considered to be a small step towards this goal.

### References

- Balzer, W./ Moulines, C. U./ Sneed, J. D.: 1987, *An architectonic for science*, Dordrecht, Kluwer.
- Bourbaki, N.: 1971, *Éléments d'histoire des mathématiques*, Paris, Hermann.
- Cartwright, N.: 1983, *How the laws of physics lie*, Oxford, Clarendon Press.
- Cassirer, E.: 1907, "Kant und die moderne Mathematik", *Kantstudien* 12, 1-49.
- Cassirer, E.: 1944, "The concept of group and the theory of perception", *Philosophy and Phenomenological Research* 5, 1-36.
- Cassirer, E.: 1980 (1910), *Substanzbegriff und Funktionsbegriff*, Darmstadt, Wissenschaftliche Buchgesellschaft.
- Cassirer, E.: 1985 (1923), *Philosophie der symbolischen Formen*, Darmstadt, Wissenschaftliche Buchgesellschaft.
- Churchland, P. M./ Hooker, C. H. (Eds.), *Images of science*, Chicago, The Univ. of Chicago Press.
- Dieudonné, J.: 1976, "Le développement historique de la notion de groupe", *Bulletin de la Société Mathématique de Belgique* 28, 267-296.
- Dornhoff, L.: 1971, *Group representation theory* (two parts), New York, Dekker.

<sup>18</sup> The analogue claim for the philosophy of empirical sciences has been widely recognized. We plea that it should be recognized in the case of mathematics too.

<sup>19</sup> Cf. (Churchland / Hooker 1985).

- Duhem, P.: 1954 (1908), *The aim and structure of physical theory*, Princeton, Princeton UP.
- Giere, R.: 1988, *Explaining science*, Chicago, University of Chicago Press.
- Goldblatt, R.: 1979, *Topoi - The categorial analysis of logic*, Amsterdam, North Holland.
- Goldman, A. I.: 1986, *Epistemology and cognition*, Cambridge, Mass., Harvard UP.
- Goodman, N.: 1978, *Ways of worldmaking*, Indianapolis, Hackett.
- Goodman, N./ Elgin, C. Z.: 1988, *Reconceptions in philosophy and other arts and sciences*, Indianapolis, Hackett.
- Hallett, M.: 1979, "Towards a theory of mathematical research programmes I, II", *British Journal for Philosophy of Science* 30, 1-25, 135-59.
- Howson, C.: 1979, "Methodology in non-empirical disciplines". In: G. Radnitzky, G. Anderson (Eds.) *The structure and development of science*, Dordrecht, Reidel, 257-266.
- Lakatos, I.: 1976, *Proofs and refutations, the logic of mathematical discovery*, Cambridge, Cambridge UP.
- Lakatos, I.: 1978, *Mathematics, science and epistemology*, Cambridge, Cambridge UP.
- Lawvere, F. W.: 1969, "Adjointness in foundations", *Dialectica* 23, 281-296.
- Mac Lane, S.: 1986, *Mathematics, form and function*, N. York-Berlin, Springer.
- Margenau, H.: 1935, "Methodology of modern physics" (two parts), *Philosophy of Science* 2, 48-72 and 164-187.
- Mundy, B.: 1986, "On the general theory of meaningful representation", *Synthese* 67, 391-437.
- Nagel, E.: 1961, *The structure of science*, London, Routledge and Kegan Paul.
- Neisser, U.: 1976, *Cognition and reality*, San Francisco, Freeman.
- Peirce, C. S.: 1931/35, *Collected papers of Charles Sanders Peirce*, six volumes, (Eds. Ch. Hartshorne, P. Weiss), Cambridge, Mass., Belknap Press.
- Putnam, H.: 1979, "Mathematics without foundations". In: *Mathematics, matter, and method*, Cambridge, Mass., Cambridge UP, 43-79.
- Quine, W. V. O.: 1979, *Philosophy of logic*, Englewood Cliffs, Prentice Hall.
- Quine, W. V. O.: 1986, "Replies to his critics". In: L. E. Hahn, P. A. Schilpp (Eds.), *The Philosophy of W. V. O. Quine*, La Salle, Ill., Open Court.
- Resnik, M. D.: 1988, "Mathematics from the structural point of view", *Revue Internationale de Philosophie* 42, 400-424.
- Suppe, F. (Ed.): 1974, *The structure of scientific theories*, Urbana, University of Illinois Press.
- Suppes, P.: 1989, "Representation theory and the analysis of structure", *Philosophia Naturalis* 25, 254-268.
- van Fraassen, B.: 1980, *The scientific image*, Oxford, Clarendon Press.
- Wussing, H.: 1969, *Die Genesis des abstrakten Gruppenbegriffs*, Berlin, VEB Deutscher Verlag der Wissenschaften.
- Yuxin, Z.: 1990, "From the logic of mathematical discovery to the methodology of scientific research programmes", *British Journal for the Philosophy of Science* 41, 377-399.

# Reduction and Explanation: Science vs. Mathematics

VEIKKO RANTALA (Tampere)

## *1. Introduction*

The aim of this essay is to compare the explanatory roles which the notion of reduction has played in the philosophy of science, on the one hand, and in the philosophy of mathematics, on the other, and to argue that in that respect there is a crucial difference between the two fields of study. Thus, for instance, in the philosophy of science the notions of explanation and reduction have been extensively discussed, even in formal frameworks, but there exist few successful and exact applications of the notions to actual theories, and, furthermore, any two philosophers of science seem to think differently about the question of how the notions should be reconstructed. On the other hand, philosophers of mathematics and mathematicians have been successful in defining and applying various exact notions of reduction (or interpretation), but they have not seriously studied the questions of explanation and understanding.

There are several reasons why reduction has been extensively discussed in the philosophy of science and in science itself. For example, it is often assumed that behind an observed, or otherwise given, phenomenon there exists a more fundamental reality to which the phenomenon can be reduced and which can be employed to explain and understand it. Secondly, it is usually thought that scientific research is not feasible if it cannot be reduced to methods which in some sense are objective and reliable. Philosophy and science abound in historical examples and consequences of these ontological and methodological forms of reductionism; such are radical empiricism and rationalism, the idea that the axiomatic method is reliable (these examples represent methodological reductionism deriving from the struggle for epistemic certainty), reductive materialism and idealism, the discussion concerning the reduction of biology to physics (which, in turn, represents ontological reductionism), discoveries of elementary particles (which are a consequence of a kind of ontological reductionism), etc.

The notion has also been important in debates concerning scientific change since it is often held – or, rather, it used to be a common view in the philosophy of science before Kuhn and other critics – that one indication of scientific progress is that theories are reducible, in some accurate, approximate, or limiting sense, to their successors, so that the latter theories are more comprehensive and more advanced than the former. A reduction, in turn, was thought to imply an explanation: if a theory is reducible to its successor, or to another more comprehensive theory, it follows that the latter explains the former in (something like) the sense of deductive-nomological explanation. Since the explanation is something that increases understanding, we should, after the reduction, be in a better position to see the nature of the reduced theory, and also the nature of the change in question.

## 2. *Scientific Progress and Reduction*

To place that discussion about reduction in its proper context, let us make next a quick survey of some earlier views in the philosophy of science concerning the question of what scientific progress might mean and whether it has been progressive. By 'progress' I shall mean here progress within a given science, that is, I shall consider theories belonging to the same branch of science. It will be of some interest to compare those earlier views with more recent ones and, in particular, with views concerning progress in mathematics. An important characteristic of progress mentioned in the literature is that scientific knowledge grows *cumulatively*, which means that old theories and the knowledge they represent survive, to some extent at least, when new and better ones appear, so that the latter extend the domain of scientific knowledge. Another characteristic is that a theory is *reducible* to its successor, which means that the old theory is not supplanted by the new theory but can be thought of as being included in the new one as a special case. A third feature is that a new theory *explains* its predecessor, so that at least the most important principles, or laws, of the old theory can be deduced from those of its successor together with some auxiliary hypotheses. These characteristics go hand in hand, more or less, whereas the following two criteria are based on somewhat different ideas (but how far they are from the first three depends on how they are interpreted). One of them says that a science is progressive if new theories *solve* at least the same *problems* as, or more or better problems than, their predecessors, and the other that progress means that scientific knowledge *approaches the truth*, that is, that new theories are better than their predecessors if they are closer to the truth and hence describe and explain the world more accurately.

These are, perhaps, the main characteristics of scientific progress which one can find in the literature, where it has been maintained, moreover, that they apply to developments in 'mature' sciences, such as modern physics. Different authors defend different characteristics. For instance, many empiricist philosophers, as, e.g., philosophers close to logical empiricism or its heir, the so-called Received View, held the view that the first three features are typical of modern science (see, e.g., Suppe, 1974), whereas some of their critics emphasize the fourth or fifth feature (e.g., Popper, 1962; Laudan, 1977; see also Niiniluoto, 1984, 1987; Pearce, 1987).

One of the most outstanding representatives of the former view is Nagel (1961) whose classical concept of reduction has usually been assumed when the view has been advocated. In short, that a theory *T* is reducible to another theory *T'* means, in the formal sense, that the laws of *T* are deducible from the laws of *T'* together with appropriate auxiliary assumptions, some of which may link the languages of the two theories to each other. It follows that the reducing theory *T'* then explains the reduced theory *T* in the sense of deductive-nomological explanation, provided, for instance, that the auxiliary assumptions satisfy appropriate theoretical and pragmatic conditions of adequacy.

According to Nagel, it is an undeniable feature of modern science that theories have been reduced to more inclusive theories, and he assumes that reduction will play an important role in the future. As standard examples of reduction in the Nagelian sense it is usually mentioned that rigid body mechanics is reducible to classical particle mechanics, Kepler's laws of planetary motion to Newton's gravitational theory, classical particle mechanics to relativistic particle mechanics, and so forth.

However, Kuhn (1962), Feyerabend (1971), and other critics of the Received View have attacked Nagel's view by arguing, for instance, that the meanings of scientific terms may change when theories change, auxiliary hypotheses are often counterfactual, and a complete translation establishing a connection between the terms of the two theories is not always possible, whence there in fact exist no proper reductions in many actual cases where reductions were claimed to exist. Hence, no intertheory explanations are available in such cases. The relation of Kepler's laws and Newton's gravitational theory and that of classical particle mechanics and relativistic particle mechanics exemplify the kind of scientific change, radical change which Kuhn (1962) calls 'revolutionary', to which the Nagelian concept of reduction and explanation is not applicable. Since this holds, it has been argued, we have to reject the view that in such cases any scientific progress has taken place in the sense of cumulation, reduction, or explanation; and the criticism also seems to challenge the view that there has been progress in the sense that the problem

solving power or the truthlikeness of theories has increased (see, however, Laudan, 1977; Niiniluoto, 1987; see also Pearce, 1987).

There are other problems, which are logical. Since many radical and important changes, particularly in physics, are such that some auxiliary hypotheses that would be needed to establish reduction relations are counterfactual, and since the theories involved in a reduction can be mutually incompatible, it has been argued that in such a case the derivation needed for an alleged reduction relation either does not make sense or is only valid approximately or in the limit, whence it follows that there is no nontrivial logical connection between the two theories of the kind required by Nagel. Thus, for instance, Newton's gravitational theory is strictly speaking incompatible with Kepler's laws and relativistic particle mechanics with classical particle mechanics, whence in these cases – in order to 'derive' in each case the latter theory from the former and thus to establish a reduction – we need such counterfactual assumptions as that the forces between the planets can be neglected or thought of as being infinitesimally small and that the velocity of light approaches infinity.

These difficulties have created uncertainty concerning the logical and explanatory status of intertheory relations, but they have also led to new attempts to study reduction. Several kinds of reduction, which more or less modify Nagel's model, have been subsequently suggested in the literature, as, for example, the models of (i) counterfactual reduction (Glymour, 1970; Pearce and Rantala, 1985; Rantala, 1989, 1991), (ii) nonlinguistic reduction (Sneed, 1971; Balzer, Moulines, and Sneed, 1987), (iii) reduction as factualization (Krajewski, 1977; Nowak, 1980), and (iv) approximate reduction (Mayr, 1981; Moulines, 1981; Niiniluoto, 1987). While Nagel's reduction should give rise to deductive-nomological explanation, the explanatory import of the other models is far from being clear. As I have argued elsewhere (Rantala, 1991), it is not evident, for instance, that if a theory *T* were approximately reduced to another theory *T'*, it would provide an explanation of *T'*; for, even though it yielded an explanation of a theory *T\** which in some sense is an approximation of *T* – this is what the approximate reduction amounts to – it would not necessarily provide a conceptual relationship of *T'* and *T* of a kind which would be needed for an explanation (cf., e.g., Tuomela, 1985, however). Reductions of the other forms (i)-(iii) are problematic as well, this time since there does not seem to be any relevant why-questions to which they would provide answers. The explanatory import of the kinds of question to which these reductions provide answers is not quite evident, whence we have to ask whether they play any role when one tries to explain, or perhaps understand, reduced theories by means of reducing ones. In this paper (in the next section), I shall only recapitulate some of the main points of the problem by referring to a general notion of reduction which is applicable to all forms of reduction

indicated above. (For a more thorough discussion, see, e.g., the articles mentioned in (i), above).

### 3. Generalized Reduction and Explanation

In mathematical logic, there exist general notions of reduction establishing both syntactic and semantic relations of theories which together can be thought of as indicating meaning change. It can be shown that even on a very general but nontrivial notion of reduction a reduction relation holding between two theories implies that the reduced theory is a logical consequence of the reducing theory and additional hypotheses. But whether it can be considered as a generalized reduction, in something like the Nagelian sense, is determined by its logical properties and the pragmatic conditions one can assign to it, and, as we have seen, such properties and conditions are even more crucial if the consequence relation is to establish a deductive-nomological explanation of the reduced theory.

Actual scientific theories, or laws, and their possible reductions are, of course, expressed by means of nonformal scientific or mathematical languages, but in order to discuss reduction in explicit terms, I shall here assume that they are formally representable in appropriate logics. (For a more comprehensive discussion about this assumption and about the forthcoming definitions, see, e.g., Pearce and Rantala, 1983, 1984, 1985; Pearce, 1987). Assume, therefore, that each theory to be considered determines a class of models (of a given type) which is definable in an appropriate logic. Since there always exist several logics in which a theory can be defined (if it can be defined in one), which of them is chosen depends on methodological, pragmatic, and purely logical criteria.

Let now  $T$  and  $T'$  be two theories such that their classes of models are  $M$  and  $M'$  and types  $t$  and  $t'$ , respectively. Assume that a logic  $L$  is assigned to  $T$  and  $L'$  to  $T'$ , and that  $\text{Sent}_L(t)$  is the set of all sentences of  $L$  which are of type  $t$ , and similarly for  $\text{Sent}_{L'}(t')$ . A *correspondence* of  $T$  to  $T'$ , relative to  $\langle L, L' \rangle$  is defined as a pair of mappings  $\langle F, I \rangle$ :

$$(3.1) \quad \begin{aligned} F: K' &\rightarrow_{\text{onto}} K \\ I: \text{Sent}_{L'}(t') &\rightarrow \text{Sent}_L(t), \end{aligned}$$

where  $K \subseteq M$  and  $K' \subseteq M'$  are definable in  $L$  and  $L'$ , respectively, such that the following condition holds for all  $m \in K'$  and all  $A \in \text{Sent}_{L'}(t')$ :

$$(3.2) \quad F(m) \models_L A \text{ iff } m \models_{L'} I(A).$$

(where  $\models_L$  denotes both the truth relation of  $L$  and logical consequence in  $L$ , and similarly for  $\models_{L'}$ ).



The existence of a correspondence relation, as above, means that there is a *weak reduction* of  $T$  to  $T'$ , and if appropriate logical and pragmatic constraints hold, an explanation of some kind is forthcoming. What this explanation is depends, e.g., on the properties of  $F$  and  $I$ . Assuming, for simplicity, that  $M$ ,  $M'$ , and  $K'$  are definable by single sentences (of  $L$  and  $L'$ , respectively)  $A$ ,  $A'$ , and  $B$ , it follows that

$$(3.3) \quad A', B \models_{L'} I(A),$$

and if  $L'$  is complete, we have, instead of Nagel's formula (see Section 2):

$$(RI) \quad A', B \vdash_{L'} I(A).$$

This shows that  $T'$  *may* explain  $T$  *via* translation, that is, that it may explain a translation of  $T$  in the language of  $T'$ , not necessarily  $T$  itself in any straightforward sense. Since the questions as to what conditions are needed in order to have an explanation and what kinds of explanation might be obtainable is discussed in earlier papers, I shall only make some brief remarks on the matter. It is conceivable that if we understand the relation of the two languages established by the translation  $I$ , and hence the meaning changes the translation takes care of, then (RI) might provide a roundabout kind of explanation of  $A$ , or  $I(A)$ , rather than any simple deductive-nomological explanation. Thus, for instance, if  $B$  is considered counterfactual, i.e., false (but not incompatible with  $A'$ , however) and  $I(A)$  is considered false, then, as I have argued in my (1989), we may interpret (RI) as providing a first step of a counterfactual explanation, that is, explanation answering the question – which modifies the question suggested by Glymour (1970) – 'On what conditions would  $I(A)$  hold?'.

A clear-cut example is provided by the question 'On what conditions would Newton's second law almost hold' (where 'almost' means infinitesimal accuracy) which, assuming that an inference like (RI) and a number of other conditions hold, can be answered, e.g., by saying that 'If the velocity of light were infinite, then Newton's second law would almost hold'. This example amounts to a logical reconstruction (in the framework discussed here) of the claim that Newton's second law is the limiting case of Minkowski's force law as the velocity of light approaches infinity. The details of this example are, however, too lengthy and tedious to be presented here (but see Pearce and Rantala, 1984; Rantala, 1989, 1991). Let us only notice here that the correctness of the above answer can be justified by means of pragmatic conditions which are 'paradigmatic' in some Kuhnian sense and which, in addition, may involve subjective attitudes of the explainer. That is, there is no 'objective' answer over and above the purely logical and mathematical components of the reconstructed reduction. Logical and mathematical features of an explanation

which is obtainable from a counterfactual reduction and its paradigmatic and intentional aspects are here even more intertwined than in deductive-nomological explanations. Another, closely related, conclusion which can be derived from case studies concerning counterfactual explanation is that whether or not the answer which such an explanation yields, i.e., a counterfactual conditional of the kind mentioned above, is true is not decidable in purely extensional means but is also a matter of *interpretation*. This fact is of such a nature that it might lessen the explanatory character (in a traditional sense) of counterfactual reduction, but, on the other hand, the reduction plays an important role in attempts to *understand* the respective scientific change (cf. Rantala, 1991). It is not quite clear, however, to what extent similar conclusions would hold for the other generalized kinds of reduction listed in Section 2, above.

#### 4. *Reduction and Explanation in Mathematics*

Do similar problems occur with respect to theories of pure mathematics, i.e., so far as mathematics is considered as a nonempirical discipline? Reasons to be interested in reduction and explanation are partly the same in metamathematics as in metascience. As we saw above, for example, philosophers of science have been at pains to ask whether scientific change is progressive, and reduction has been a tool which is used to answer such questions. Philosophers of mathematics have also been interested in the problem of progress, but it seems, on the other hand, that they have not studied, so much, the role of explanation in mathematical developments.

One prominent exception is Kitcher (1983, p. 227) who distinguishes three types of mathematical explanation, which are connected with different kinds of mathematical progress. First, "...we can sometimes explain mathematical theorems by recognizing ways in which analogous results would be generated if we modified our language". This is connected with extending mathematical language by generalization. A well-known example is Cantor's generalization of finite arithmetic. Generalizations may explain by showing how a generalized language and theory are obtained within which results analogous to those we have already accepted are forthcoming. From the point of view of the generalization we can see the old theory as its special case, and at the same time the generalization may improve our understanding of the old theory (Kitcher, 1983, p. 209).

A second type of explanation is connected with a clarification of language and of techniques of reasoning, and it is called 'rigorization' by Kitcher. Obviously, one of the most dominating features in the historical developments of mathematical practice is that it has become more and more rigorous, and,

hence, examples can be easily found, for instance, in the development of analysis. Rigorization can be explanatory in that it removes "...previous inability to recognize the fine structure of connections". The third type mentioned by Kitcher is associated with what he calls 'systematization', by which he means such activities as axiomatization and conceptualization, where the latter term refers to a modification of mathematical language "...so as to reveal the similarities among results previously viewed as diverse or to show the common character of certain methods of reasoning" (Kitcher, 1983, p. 218). Systematization is explanatory in that it yields unification.

Explanations of the mentioned kinds at least satisfy the most important requirement which all explanations have to satisfy, that is, they increase our understanding, but, on the other hand, Kitcher does not make it very clear what the explanation-seeking questions would be like which these explanations are assumed to answer. Each of these types is so general that we have to look at their specific applications in order to see the corresponding questions, but it is obvious, however, that the cases of generalization, rigorization, and unification mean mathematical progress and, hence, may provide answers to why-questions of some kinds. Similar patterns of explanation may occur in scientific practice as well, but so far as the explanation of *theories* is concerned, be they mathematical or scientific, the notion of reduction – whose special case generalization obviously is – is the most central tool, as we already noticed, and, on the other hand, the notion of correspondence which we defined earlier, in Section 3, covers the most important notions of reduction used for various metamathematical purposes.

To see whether the actual cases of reduction in mathematics really have an explanatory import and how reductive explanations in mathematics would differ from those in science, we should work out detailed and comparative case studies. There exist developments in mathematics – and they seem to be less controversial than many of the much discussed changes in science – which can be considered progressive and where reduction seems to play a similar explanatory role as in science. On the other hand, however, this role is not always very obvious. It is not quite clear, for example, in what sense a reduction of arithmetic to set theory really explains arithmetics even though it may increase our understanding. As it is usually recognized, this reduction, among many other reductions in mathematics, is of ontological and methodological importance. According to Bonevac (1982), its ontological import is, e.g., due to the fact that sets are epistemologically at least as accessible as numbers – since numbers are at least as abstract as the sets with which they are 'identified', i.e., to which they are reduced. Bonevac's notion of epistemic accessibility seems to be more or less empiricist, however – in the sense that "...our ability to have knowledge concerning the objects assumed to exist

must itself be capable of being a subject for empirical, and preferably physiological, investigation" (Bonevac, 1982, pp. 8-9) – and, hence, could be criticized in the same way as the empiricists' views concerning the cognitive role of theoretical entities in science have been criticized during the last thirty years (see, Suppe, 1974).

Whether or not ontological reduction has epistemic importance of the kind advocated by Bonevac, its methodological value is undeniable at least in cases where reductions are part of systematization in the sense of Kitcher (1983). It is this methodological sense, rather than ontological or epistemic, in which, for example, the reduction of arithmetic to set theory may yield an explanation of the former theory and may advance our understanding of the role of numbers.

We can see now that even though the aims of reduction in mathematics are, at least in part, the same as in science, its explanatory roles in these two fields are in many ways different. So far as mathematical progress is concerned, there hardly exist in (pure) mathematics any important cases of counterfactual or approximate reduction and explanation, in the sense discussed in Section 3, since in all important reductions in mathematics the reduced theories are not considered false – whence it follows, in particular, that the extensional notion of deductive-nomological explanation is more readily applicable in mathematics, as can be expected. The role of pragmatic and intensional features of explanations derived from mathematical reductions is as minimal as it can be. Pragmatic features seem to occur more prominently in the kinds of explanation in which Kitcher is concerned. As we saw, counterfactual and approximate reductions are, on the other hand, the most important kinds of reduction that are associated with actual theories in science.

It is argued by Bonevac (1982, p. 55) that in mathematics there exist no counterfactual reductions since in mathematical contexts the notion of counterfactual does not make much sense. It does not make sense, according to Bonevac, since mathematical statements are necessarily true if they are true at all. This is a bit hasty conclusion, however, since it is easy to find mathematical statements which are not necessary nor false in any straightforward sense. Thus, for instance, the statement, formulated in an appropriate language, claiming that there exists a number which is greater than all natural numbers is counterfactual as far as standard models of arithmetic are concerned but true in its nonstandard models – and there are, of course, indefinitely many similar examples, all associated with generalizations. Whether or not a statement is necessary is, even in mathematics, more or less relative to the context and logic within which it is considered, and hence the reason why there exist no important counterfactual reductions in pure mathematics is not that the notion of counterfactual does not make sense.

### 5. *Change in Logic*

There is another kind of theoretical change which has been considered more relevant for mathematics than for empirical science and which is similar to, but in many ways even more fundamental than, paradigm change in science. It consists of changing the logical foundations of theories. The role of explanation is here even more problematic since it is not always clear how far (for instance, in metatheory) such a change – in particular, if it is radical – would eventually go or should go.

Furthermore, philosophical discussions about logical change have been ambiguous, so far, in the sense that no clear distinction has been made between the logic of a given theory (i.e., the logic in which the theory is reconstructed, if in any, or whose principles the inferences of the theory and its interpretations are assumed more or less to follow) and the logic of its metatheory (see, e.g., Quine, 1970; Haack, 1974; Briskman, 1982). That is to say, it is difficult to see what role logic is assumed to play here. The latter logic is usually even more ambiguous than the former, and, as we know well, there is considerable disagreement concerning the question what the logic is that underlies mathematical reasoning or natural language, that is, underlies the metatheory of possibly formalized theories – if there underlies any definite logic – and what it possibly should be. In what follows, I will nevertheless assume, to be able to argue in explicit terms, that to each theory some logic is assigned in which the theory is formalized – and which may change – and that the metatheory, even though not formalized, is dependent on some recognizable logical principles – which may also change.

So, instead of thinking about the change of theories in the first place, let us emphasize now the change of logics. This emphasis is more or less hypothetical since, for evident reasons, there exist no obvious cases of revisions of logics underlying actual empirical theories: there usually are no definite, explicitly characterized logics for actual theories, not before they are logically reconstructed, and, furthermore, scientists themselves are not interested in such questions. On the other hand, philosophers of science have suggested revisions, as for instance, in the case of quantum logic where the suggestion is a consequence of a new interpretation of 'empirical reality'. Purely logical reasons for revisions have been suggested as well, which, however, have so far concerned mathematical theories rather than empirical ones and, hence, resulted in proposals for adjusting mathematical principles. The cases of intuitionistic logic and relevance logic provide examples.

Keeping to the aims of the present essay, I shall only consider the kind of change where an earlier logic is in some sense reducible to the new one. What would then be a weakest notion of reduction such that it would cover different

variants of reduction and still make it possible to say that the reduced logic is explained by the reducing one? I shall briefly study one suggestion (Rantala, 1988), whose explanatory import depends, again, on context and on the pragmatic constraints the mappings involved satisfy and which leads to a straightforward modification of the notion of correspondence defined earlier in Section 3.

Since for our purposes here it is not necessary to study the general notion of a logic in greater detail, let us only think of a logic as something having both syntactic and semantic components which satisfy appropriate conditions, most of which pertain to the idea that logics are considered extensional. More precisely, by a logic I mean here what it means in general model theory (see Feferman, 1974, whose terminology and notation will be used below). To each logic  $L$ , the following components are assigned: (i) the class of all similarity types admitted by  $L$ ,  $\text{Typ}_L$ ; and for each type  $t \in \text{Typ}_L$ , (ii) the class of all admitted models of type  $t$ ,  $\text{Str}_L(t)$ , (iii) the class of all  $L$ -sentences of type  $t$ ,  $\text{Sent}_L(t)$ , and (iv) a truth relation  $\models_{L(t)}$  (in what follows, ' $\models_L$ ', for short).

The following definition seems to generalize various existing accounts in logic, and at the same time it is analogous to the earlier definition of correspondence. A logic  $L$  is *reducible* to a logic  $L'$  if for all  $t \in \text{Typ}_L$ , there exist a type  $t' \in \text{Typ}_{L'}$ , an  $L'$ -definable class  $M \subseteq \text{Str}_{L'}(t')$ , and mappings

$$(5.1) \quad \begin{aligned} F: M &\rightarrow_{\text{onto}} \text{Str}_L(t), \\ I: \text{Sent}_L(t) &\rightarrow \text{Sent}_{L'}(t') \end{aligned}$$

such that the following holds for all  $m \in M$ ,  $A \in \text{Sent}_L(t)$ :

$$(5.2) \quad F(m) \models_L A \text{ iff } m \models_{L'} I(A).$$

If  $L$  is reducible to  $L'$ , important properties transfer from  $L'$  to  $L$ , such as Compactness and Löwenheim. Since the translations of the valid sentences of  $L$  are logical consequences in  $L'$  of the sentences defining  $M$ ,  $L'$  may explain  $L$  – on appropriate pragmatic conditions – in some of the senses which are analogous to those discussed earlier in connection with correspondence. Furthermore, any theory  $T$ , mathematical or scientific, formulated in  $L$  can be reformulated in  $L'$  in the obvious way, that is, if  $T$  is axiomatizable in  $L$ , it can be translated into a theory  $T'$  in  $L'$ , whence there is a correspondence of  $T$  to  $T'$ , relative to  $\langle L, L' \rangle$ .

It can be shown, for instance, that classical predicate logic is reducible to intuitionistic one in the above sense, and *vice versa*, but we have to ask, again, whether, and in what sense, these reductions are explanatory. Therefore, they should be studied more closely from a pragmatic point of view; but it is obvious, anyway, that when classical logic is reduced to intuitionistic logic,

the character of a possible explanation depends on the philosophical role the Kripke semantics of intuitionistic logic is assumed to play. Thus, for instance, if the semantics is given the epistemic construal, as it is often done, the reduction may explain the status of classical logic from the epistemic point of view which is represented by intuitionistic logic. In any case, the reduction increases our understanding of the relation of the two logics – whether they are considered from a formal or epistemic point of view. Furthermore, since there is a correspondence of a theory  $T$  axiomatizable in classical logic to its translation in intuitionistic logic,  $T'$ , it immediately follows that if the former logic is in some sense explained by the latter, or understood in terms of the latter, the corresponding fact holds for  $T$  and  $T'$  as well.

## 6. Conclusion

In the philosophy of science, much attention has been paid to the notion of explanation, but no agreement upon its form or import has resulted from this interest. In the philosophy of the human sciences and in aesthetics, understanding and interpretation have been central notions, as well, and they have appeared even more controversial than explanation. In the philosophy of mathematics, on the other hand, more explicit attention should perhaps be paid to these and related notions, so that we would be in a better position to understand the cognitive and symbolic nature of mathematical change. As we have seen, for instance, it is far from clear what the cognitive import of reduction basically is in mathematics – in particular, that of ontological reduction, over and above the more or less empiricist views which have been dominating much of recent philosophy of mathematics. More importantly, the question concerning the nature of mathematical progress is to a great extent a matter of interpretation.

## References

- Balzer W. / Moulines, C. U. / Sneed, J. D. (1987), *An Architectonic for Science*, D. Reidel, Dordrecht.
- Bonevac, D. A. (1982), *Reduction in the Abstract Sciences*, Hackett Publishing Company, Indianapolis-Cambridge.
- Briskman, L., "From Logic to Logics (and Back Again)", *The British Journal for the Philosophy of Science* 33 (1982), 95–111.
- Feyerman, S., "Two Notes on Abstract Model Theory I", *Fundamenta Mathematica* 82 (1974), 153–165.
- Feyerabend, P. (1975), *Against Method*, NLB, London.
- Glymour, C., "On Some Patterns of Reduction", *Philosophy of Science* 37 (1970), 340–353.
- Haack, S. (1974), *Deviant Logic*, Cambridge UP, Cambridge.

- Kitcher, P. (1983), *The Nature of Mathematical Knowledge*, Oxford UP, New York-Oxford.
- Kuhn, T. (1962), *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago.
- Krajewski, W. (1977), *Correspondence Principle and the Growth of Knowledge*, D. Reidel, Dordrecht.
- Laudan, L. (1977), *Progress and Its Problems*, University of California Press, Berkeley, Los Angeles-London.
- Mayr, D. (1981), "Approximate Reduction by Completion of Empirical Uniformities". In: A. Hartkämer and H.-J. Schmidt (eds.), *Structure and Approximation in Physical Theories*, Plenum Press, New York, 55–70.
- Moulines, C. U. (1981), "A General Scheme for Intertheoretic Approximation". In: A. Hartkämer and H.-J. Schmidt (eds.), *Structure and Approximation in Physical Theories*, Plenum Press, New York, 123–146.
- Nagel, E. (1961), *The Structure of Science*, Routledge & Kegan Paul, London.
- Niiniluoto, I. (1984), *Is Science Progressive?*, D. Reidel, Dordrecht.
- Niiniluoto, I. (1987), *Truthlikeness*, D. Reidel, Dordrecht.
- Nowak, L. (1980), *The Structure of Idealization*, D. Reidel, Dordrecht.
- Pearce, D. (1987), *Roads to Commensurability*, D. Reidel, Dordrecht.
- Pearce, D. / Rantala, V., "New Foundations for Metascience", *Synthese* 56 (1983), 1–26.
- Pearce, D. / Rantala, V., "A Logical Study of the Correspondence Relation", *Journal of Philosophical Logic* 13 (1984), 47–84.
- Pearce, D. / Rantala, V., "Approximate Explanation Is Deductive-Nomological", *Philosophy of Science* 52 (1985), 126–140.
- Popper, K.R. (1962), "Some Comments on Truth and the Growth of Knowledge". In: E. Nagel, P. Suppes, and A. Tarski (eds.), *Logic, Methodology, and Philosophy of Science: Proceedings of the 1960 International Congress*, Stanford UP, Stanford, 285–292.
- Putnam, H. (1975), "The Logic of Quantum Mechanics". In: H. Putnam, *Mathematics, Matter and Method*, Cambridge UP, Cambridge, 174–197.
- Quine, W. V. O. (1970), *Philosophy of Logic*, Prentice-Hall, New Jersey.
- Rantala, V. (1988), "Scientific Change and Change of Logic". In: I. M. Bodnar, A. Maté, and L. Polos (eds.), *Intensional Logic, History of Philosophy, and Methodology*, Budapest, 247–252.
- Rantala, V. (1989), "Counterfactual Reduction". In: K. Gavroglu, Y. Goudaroulis, and P. Nicolacopoulos (eds.), *Imre Lakatos and Theories of Scientific Change*, Kluwer Academic Publishers, Dordrecht, 347–360.
- Rantala, V. (1991), "Understanding Scientific Change". In: P. Bystron and V. Sadovsky (eds.), *Festschrift in Honour of Vladimir A. Smirnov*, Kluwer Academic Publishers, Dordrecht, forthcoming.
- Sneed, J. D. (1971), *The Logical Structure of Mathematical Physics*, D. Reidel, Dordrecht.
- Stegmüller, W. (1976), *The Structure and Dynamics of Theories*, Springer-Verlag, Berlin-Heidelberg-New York.
- Suppe, F. (1974), *The Structure of Theories*, University of Illinois Press, Urbana, Illinois.
- Tuomela, R. (1985), *Science, Action, and Reality*, D. Reidel, Dordrecht.



## Reality, Truth, and Confirmation in Mathematics – Reflections on the Quasi-Empiricist Programme

ILKKA NIINILUOTO (Helsinki)

### 1. *New Trends in the Philosophy of Mathematics*

Ever since the days of Plato, the philosophy of mathematics has been dominated by the view that mathematical truth is *a priori*, independent of sense experience. According to the platonist, a mathematician uses his light of reason, or non-sensuous intuition, to uncover eternal and necessary truths about a pre-existing domain of abstract objects. Immanuel Kant's doctrine claimed that both arithmetic and geometry are synthetic *a priori*. During our century, the main schools have explained the special character of mathematical knowledge by claiming that a mathematician studies mental constructions in his own mind (intuitionism), games of manipulating syntactical signs (formalism), or logical tautologies without any factual content (logicism).<sup>1</sup>

None of these rival approaches takes seriously the empiricist thesis, defended by John Stuart Mill in his *A System of Logic* (1843), that the basic truths of arithmetic and geometry are inductive generalizations from experience. Gottlob Frege's ironic and devastating criticism, in *Die Grundlagen der Arithmetik* (1884), have been taken as a conclusive refutation of Mill: the empiricist philosophy of mathematics confuses pure deductive mathematics, based upon proof, and the applications of mathematics. In this spirit, the logical empiricists argued that pure mathematics (e.g., axiomatic geometry) is *a priori* and analytic, while mathematics applied to reality (e.g., physical geometry) is a branch of natural science.<sup>2</sup>

---

<sup>1</sup> These positions are well represented in Benacerraf / Putnam (1964) with articles by Brouwer, Hilbert, Frege, and Russell.

<sup>2</sup> A classical expression of this view of geometry is Reichenbach (1957) (appeared originally 1928). For the distinction between analytic and synthetic truth, see Frege (1959) and Stenius (1972). For the possibility of making this distinction within empirical theories, see Tuomela (1973).

While the formalists, the logicians, and the intuitionists have relied primarily on syntactical or proof-theoretical methods in their foundational studies, model theory pictures mathematics as a study of structures. This view, common to Alfred Tarski and Bourbaki's "structuralism", again usually tends towards some form of platonism, since mathematical structures can be regarded as set-theoretical entities "living" in an abstract universe of sets. Some philosophers would assume that there is a unique standard model for set theory, others assert the existence of several alternative universes – but, in any case, knowledge about them is founded on proof or *a priori* reasoning.<sup>3</sup>

Recent developments in the philosophy of mathematics have raised interesting and important challenges to these dominant approaches. One of these new trends is concerned with the actual methodology of mathematics. George Polya suggested already in 1945 that "mathematics presented with rigor is a systematic deductive science but mathematics in the making is an experimental inductive science".<sup>4</sup> Imre Lakatos distinguished "Euclidean" and "quasi-empiricist" mathematics, claiming that the latter follows the Popperian method of conjectures and refutations.<sup>5</sup> The successful use of computers in the proof of mathematical theorems, such as the four-colour theorem, has given further strength to the quasi-empiricist approach – as witnessed by Thomas Tymoczko's excellent collection *New Directions in the Philosophy of Mathematics* (1986).

Another trend is the recovery of the Millian theory of numbers as properties of aggregates. Peter M. Simons (1982) defends this view by reference to Husserl's theory of manifolds, and John Bigelow (1988) interprets natural numbers as relational universals (in the sense of Armstrong's physicalism). Return to Mill, with full endorsement of the empiricist doctrine that experience is the origin and ultimate foundation of mathematical knowledge, is advocated by Philip Kitcher's *The Nature of Mathematical Knowledge* (1983).

Several philosophers have recently investigated mathematics from the viewpoint of physicalist ontology. Hartry Field's *Science without Numbers* (1980) argues that numbers are dispensable and eliminable in physical theories and, therefore, can be regarded as mere fictions. Starting from a causal theory of knowledge, Paul Benacerraf and others have pointed out that causal interaction with – and, hence, knowledge about – platonic objects is impossible.<sup>6</sup>

---

3 See the articles on set theory in Benacerraf / Putnam (1983). For the model-theoretic concept of truth, see Niiniluoto (1987).

4 See Polya (1945, 1954).

5 See Lakatos (1976) and Tymoczko (1986). Cf. also Putnam's (1975) evaluation of quasi-empiricism.

6 See Benacerraf's article "Mathematical Truth" (1973), reprinted in Benacerraf / Putnam (1983).

This line of thought has led to a revival of physicalist approaches where mathematical theories are interpreted as claims about actual or possible aspects of physical reality or mathematical practice.<sup>7</sup> Many versions of physicalism go together with the contention that sense experience plays an important role in the formation of mathematical knowledge.

In this paper, I try to critically evaluate some of the main ideas of the new physicalist and quasi-empiricist trends. Starting from Karl Popper's ontological distinction between Worlds 1, 2, and 3, I suggest that it is possible to be a realist and a constructivist at the same time. Pure mathematics studies man-made structures in World 3 of abstract artefacts. Induction may have an important role in the confirmation of mathematical conjectures, but it is secondary with respect to deductive proof. While the main content of mathematical theories is about World 3, quantitative statements have – via theories of measurement – applications in the physical World 1 and the mental World 2.

## 2. *Poor Man's Platonism or How To Be a Realist and a Constructivist at the Same Time*

Traditional ontologies have accepted three kinds of entities. First, physical objects, things, and processes; secondly, mental or psychical states and events; thirdly, abstract objects, like transcendent universals, concepts, propositions, objective spirit, God, etc. Following Popper (1972), we may use *World 1*, *World 2*, and *World 3* as convenient labels for these three realms.<sup>8</sup> This gives us a nice classification of three monistic metaphysical doctrines: *materialism* (*physicalism*) claims that everything that is real belongs, or is reducible, to World 1; *subjective idealism* (*mentalism*) makes the same claim relative to World 2; *objective idealism* (*platonism*) urges that World 3 is the primary basis or source of all being. On the other hand, dualist ontologies accept World 1 and World 2 (or World 1 and World 3) as two independent domains of reality. (See Fig. 1).

---

<sup>7</sup> See Putnam's "Mathematics without Foundations" (1967), reprinted in Benacerraf / Putnam (1983), Irvine (1990), and Kitcher (1983).

<sup>8</sup> See also Niiniluoto (1984), Ch. 9.

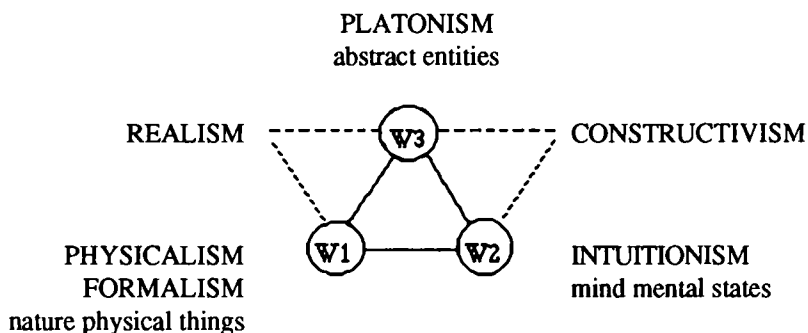


Fig. 1

Philosophies of mathematics can also be classified by our simple scheme. A *platonist* postulates that mathematical objects (numbers, geometrical figures, sets, etc.) belong to World 3; an *intuitionist* finds their place among mental constructions or World 2; a *physicalist* locates them in World 1; in particular, a formalist regards mathematics as a play with material signs (e.g., numerals).

Many moves in philosophical debates can be explained by the tacit assumption that Fig. 1 covers *all* the relevant alternatives. For example, J.S. Mill – after rejecting both formalism ("the propositions of the science of number are not verbal") and platonism (the truths of mathematics do not "relate to, and express the properties of, purely imaginary objects") – is puzzled by the fact that "points without magnitude", "lines without breadth" or "perfect squares" exist "neither in nature nor in the human mind".<sup>9</sup>

Similarly, it is often assumed that a *realist*, who wishes to defend the objectivity of mathematical truths, has to be either a platonist or a physicalist. Inference from the combination of realism and anti-platonism to a physicalist interpretation of mathematical truth is indeed a quite common theme in contemporary philosophy.<sup>10</sup> Further, it is also often assumed that a *constructivist*, who regards mathematical objects as constructs rather than as pre-existing objects, has to be either a subjective mentalist or some kind of objective idealist (e.g., mathematics as the work of a "creative subject").<sup>11</sup>

In my view, these dichotomies and inferences are unwarranted. They ignore the possibility that one can be a realist and a constructivist at the same time.

<sup>9</sup> See Mill (1906), pp. 147–148.

<sup>10</sup> See the papers in Irvine (1990).

<sup>11</sup> Brouwer's theory of the creative subject is discussed by Troelstra (1969).

This view has been defended by Popper (1972): World 3 is man-made, it "originates as a product of human activity", yet it "transcends its makers". Even though human creations, World 3 entities are real or autonomous in the sense that they interact causally with Worlds 2 and 1. World 3 "has grown far beyond the grasp, not only of any man, but even of all men (as shown by the existence of insoluble problems)". In particular, "the natural numbers are the work of men, the product of human language and of human thought". Still, there is an infinity of such numbers and true equations between them – "more than will ever be pronounced by men, or used by computers".

Popper's ontological position can be regarded as poor man's platonism: as no god has created abstract entities for us, we have to make them for ourselves! World 3, the realm of human creations, contains all cultural entities, artefacts, works of art and literature, institutions, theories, abstractions. Even if cultural objects are social constructions, they are not entirely transparent to us (as the Cartesian tradition would mistakenly assume). We can, in an objective sense, discover new truths about our own creations.<sup>12</sup>

The position that emerges from these considerations is realist and constructivist at the same time. Mathematics studies structures – like the set of natural numbers  $\mathbb{N}$ , real numbers  $\mathbb{R}$ , Euclidean space  $\mathbb{R}^n$ , etc. – which are constructed as World 3 entities by giving a finite description of the rules of their formation. These descriptions (embedded in a suitable mathematical background theory) are sufficient to guarantee that the statements of the relevant mathematical language have a determinate truth value, true or false in Tarski's sense, within these structures. All this holds in spite of the fact that an infinite set like  $\mathbb{N}$  or  $\mathbb{R}$  will never be actually realized in World 1 (e.g., by paper and pencil, or by a computer) or in World 2 (by thoughts or ideas in the human mind).<sup>13</sup>

---

<sup>12</sup> Popper's way of formulating his theory of World 3 is not unproblematic, but I believe it can be made sound and coherent in terms of emergent materialism (cf. Niiniluoto, 1984). Popper's philosophy of mathematics is discussed critically by O'Hear (1980). See also Gillies (1990). Most of the recent writers in this field fail to mention Popper at all.

<sup>13</sup> My claim here is that, even though Worlds 1 and 2 are finite, it is possible to construct infinite classes in World 3 – without having all of their elements "embodied" at the same time. In this sense, my version of "constructivism" is not "Aristotelian" (cf. Gillies, 1990). On the other hand, it is not "Platonist" either, since a World 3 entity loses its existence, if its documentation and manifestation in Worlds 1 and 2 discontinues. It is not possible here to go into details about the important question of the conditions and limits of acceptable "constructions" in mathematics.

For example, in 1989, a group of mathematicians used the Amdahl 1200 supercomputer to find the largest prime number known so far,  $391581 \times 2^{216913} - 1$ . It is probable that no one had ever thought about this particular number or written down its 65087 digits. Still, this natural number was "real" in the Peircean sense that it had (even before 1989) the property of being prime "objectively", that is, independently of what anybody may think about its characters.

It may be objected that, e.g., the set  $\mathbb{N}$  of natural numbers can be "constructed" as a set-theoretical entity in many alternative ways which are all isomorphic with each other. Paul Benacerraf concludes that arithmetic is "not a science concerned with particular objects – the numbers", but it rather "elaborates the abstract structure that all progressions have in common merely in virtue of being progressions".<sup>14</sup> This view – called "eliminative structuralism" by Parsons (1990) – attempts to get rid of reference to mathematical entities. However, an alternative conclusion from Benacerraf's argument is to say that numbers are *World 3 objects with only relational properties* (odd = not divisible by 2; etc.). Thus, World 3 objects, as man-made abstractions, can be well-defined relative to their relational arithmetic properties but indefinite relative to other properties.<sup>15</sup>

For example, number 2 can be represented set-theoretically by  $\{\emptyset, \{\emptyset\}\}$  or by  $\{\{\emptyset\}\}$ , but only some features of these constructions are relevant to the arithmetic properties of 2. World 3 has a hierarchical structure, where some objects are more abstract than their more concrete realizations.<sup>16</sup> In this sense, an abstract entity like number 2 resembles a musical composition which is a unique artefact created by a composer. For example, the 7th symphony of Sibelius can be realized in many different forms in Worlds 1 (note scripts, records, tapes, waves) or World 2 (thoughts, experiences), but still it is a unique object in World 3.

This way of thinking about mathematical objects allows us to say that several early cultures used different names (i.e., numerals) to refer to the *same*

---

<sup>14</sup> See Benacerraf / Putnam (1983), p. 291. A "modal-structural" version of eliminativism is presented by Hellman (1989), whose work has been inspired by Dedekind and Putnam.

<sup>15</sup> This incompleteness of numbers agrees with Michael Resnik's account of numbers as "positions" in "patterns". The main difference to Resnik is that his "positions" are Platonic entities, my World 3 objects are human constructions.

<sup>16</sup> Natural numbers, as unique World 3 entities, are related to their particular set-theoretic representations by what Tait (1986), p. 369, calls "Dedekind abstraction". Our account of World 3 explains also Hellman's (1989), p. 14, puzzle: how can Dedekind appear to be a platonist and a creationist as well?

natural numbers. Thus, there was a time when numbers were first invented. The same numbers have later been reinvented again and again, and their objective properties can be studied through one of the mutually isomorphic set-theoretical structures representing numbers.<sup>17</sup>

An excellent formulation – without reference to Popper – of this view is given by Ruben Hersh:

- "(1) Mathematical objects are invented or created by humans.
- (2) They are created, not arbitrarily, but arise from the activity, with already existing mathematical objects, and from the needs of science and daily life.
- (3) Once created, mathematical objects have properties which are well-determined, which we may have great difficulty in discovering, but which are possessed independently of our knowledge of them".

But even though Hersh's article appears in Tymoczko's (1986) quasi-empiricist collection, the view of mathematics as a cultural science allows only a restricted and secondary role for empirical procedures in mathematical inquiry. At least, this is what we shall argue in the next sections.

### 3. *Our Knowledge of World 3*

Our knowledge of World 3 is normally based upon our interaction with human constructions in the sphere of language, culture, and society – and thus *a posteriori*.<sup>18</sup> However, mathematical truth is analytic – and typically known *a priori*. A causal theory of knowledge, which has been used as a premise supporting physicalist-empiricist interpretation of mathematics, is plausible for our knowledge of World 1. But, we shall see, it is neither needed nor appropriate for all abstract World 3 objects.

Mathematics studies structures which belong to the man-made World 3. They are constructed by definitions which can be expressed in set theory. When *M* is a set-theoretical structure (i.e., a domain *D* of mathematical objects together with relations and functions on *D*), and *L* is a language interpreted in *M*, the sentences of *L* have truth values in *M* (in the sense of Tarski's recursive definition).

---

17 This view comes close to Michael Dummett's (1964) "intermediate picture" between platonism and constructivism. According to Dummett, mathematical objects "spring into being in response to our probing". "We do not *make* the objects but must accept them as we find them". The difference between Dummett's objects and World 3 entities seems to be that the latter can continue their public existence and can be reinvented after their creation.

18 Cf. Niiniluoto (1981).

Thus, mathematical truth is, in a semantic sense, *analytic*: if sentence  $h$  is true in structure  $M$ , its truth is a consequence of the definition of  $M$  – and independent of factual assumptions.<sup>19</sup>

But how do we learn and justify mathematical truths? To answer this epistemological question, let us note first that typical examples of mathematical knowledge include three types of statements:

- (1) Singular and existential statements about a particular structure  $M$  (e.g., '211 is a prime number', 'There is a subset of reals  $\mathbb{R}$  which is non-denumerable and nowhere dense').
- (2) General statements about a particular structure  $M$  (e.g., 'For all natural numbers  $n \in \mathbb{N}$ :  $1^2 + 2^2 + \dots + n^2 = n(n+1)(2n+1)/6$ ').
- (3) Generalizations over types of structures (e.g., 'In a distributive lattice the complement of an element is unique', 'Every metric space is a Hausdorff space', 'Every subgroup of a cyclic group is cyclic').

Perhaps the most common of them, at least in advanced mathematics, is the third case.

It is possible to obtain *a priori* knowledge of these three types of examples by using *deductive proofs*.

In case (3), let  $A$  and  $B$  be sets of sentences describing types of mathematical structures such that  $A \vdash B$  (i.e.,  $B$  is logically deducible from  $A$ ). As

- (4) If  $M \models A$  and  $A \vdash B$ , then  $M \models B$ ,

it follows from  $A \vdash B$  that

- (5)  $\forall M$  (if  $M \models A$  then  $M \models B$ ).

For example,  $A$  may be the definition of a metric space and  $B$  the definition of a topological Hausdorff space.

In case (2), the proof of the arithmetical equation uses the principle of mathematical induction. An existential statement is proved by showing how a mathematical object with the desired properties can be constructed. (In classical mathematics, also indirect proofs of existence are accepted).

The method of proof can be applied also with respect to a single structure – such as the set  $\mathbb{N}$  of natural numbers with the usual arithmetical operations.

---

<sup>19</sup> If a metalinguistic, rather than a set-theoretic, definition of structure is used (cf. Parsons, 1990), this conclusion can be expressed as follows. A sentence in language  $L$  is analytically true in  $L$  if and only if, according to the linguistic conventions  $C$  about  $L$ , sentence  $h$  is true no matter what state of affairs obtains (Stenius, 1972, p. 82). In the case of arithmetic, such conventions  $C$  serve to define the structure  $\mathbb{N}$  of natural numbers.



This is what Lakatos calls "Euclidean" mathematics: choose as axioms a set of sentences true in  $\mathbb{N}$ , e.g., the axioms PA of Peano arithmetic. Then, by (4), everything that follows deductively from PA is also true in  $\mathbb{N}$ :

(6) if  $PA \vdash h$ , then  $\mathbb{N} \models h$ .

To prove the truth of  $h$ , it is thus sufficient to deduce it from the axioms PA.

By Gödel's Incompleteness Theorem, we know, however, that the set  $Cn(PA)$  of all deductive consequences of PA is only a proper subset of the complete arithmetic  $Th(\mathbb{N})$  which consists of all arithmetical sentences true in  $\mathbb{N}$ .

In practice, a mathematician usually relies on conceptual operations and procedures which can in principle be expressed, e.g., in formal arithmetic. Even if these operations need not be infallible – mistakes in mental calculation are possible – they give us *a priori* knowledge. They may be assisted by paper and pencil, but nevertheless they are independent of evidence by sense perception.<sup>20</sup> This is the case, since these conceptual operations give us "maker's knowledge" of our own constructions. The World 3 ontology helps us to understand this important issue.

Popper suggests that our "grasping" of a World 3 object means the "making" or the "re-creation" of it.<sup>21</sup> This is not at all plausible as a general account of our knowledge about World 3, since we are not able to "re-create" such entities as paintings, novels, symphonies, societies, or legal orders. But for mathematics this view fits quite well: mathematical objects, or sets representing them, can be reproduced, whenever we wish to study their properties and relations. For example, to solve some cognitive problem about  $\mathbb{N}$  (e.g., is 211 a prime number or not?), we construct or re-create  $\mathbb{N}$  to the extent that is necessary for finding the answer.

Hence, it is not necessary that the justification of *our* mathematical knowledge is traced "through a chain of prior authorities", reaching eventually "perceptual knowledge acquired by our remote ancestors". Such a genetic account – suggested by Kitcher (1983) in his empiricist philosophy of math-

---

20 Mental calculations, which follow the recursive definition of arithmetic operations, give us an *a priori* warrant for believing numerical equations (cf. Kitcher, 1983). These mental operations may co-exist with empirical calculations with pebbles or pocket calculators (see Section 6 below). I disagree with Stenius (1972), who argues that numerical truths are synthetic and *a posteriori*, since they can be correctly verified by one (and only one) calculation. I reject a premise that Stenius is using in his argument: "a statement which is testable by empirical observation cannot also be known to be true *a priori*" (*ibid.*, p. 83).

21 See Popper / Eccles (1977), p. 44.

ematics – may be possible for some parts of mathematics (such as arithmetic or plane geometry), but not for all of its branches (e.g., what would be the perceptual origin of Hamilton's quaternions or topological Lindelöf spaces?). But even in the case of natural numbers  $\mathbb{N}$ , the justification of our arithmetic knowledge need not refer to the old days when the elements of  $\mathbb{N}$  were first created, since we can always *reproduce* them *now* – and study them for our own purposes.

#### 4. *Abduction and Induction in Pure Mathematics*

The deductivist and apriorist account of mathematical knowledge, outlined above, is compatible with the observation of the quasi-empiricist school that non-demonstrative inferences *sometimes* play an important role in mathematics. Polya illustrated this with a great number of examples, where mathematical theorems are discovered by inductive generalization or analogical inference. But their use is not restricted to "mathematics in the making", or to the context of discovery, since – as Lakatos pointed out – often the best justification we have for a mathematical conjecture arises from attempts to refute it. Thus, mathematical claims can be confirmed by the hypothetico-deductive method.

A typical example of non-demonstrative inference is the search for generality in mathematics.<sup>22</sup> If a theorem holds in  $\mathbb{R}$ , it immediately invites us to conjecture that its generalized version holds in  $\mathbb{R}^n$  for all  $n$ . This heuristic inference is an instance of Peircean *abduction*. The general claim may then be tested in the special cases for  $\mathbb{R}^2$  and  $\mathbb{R}^3$ ; a proof of these instances then gives *confirmation* to the general theorem. If the test fails, a hidden assumption of the general theorem may be found.

Typical examples of analogy include cases, where we conjecture that a theorem valid for  $\mathbb{R}$  holds also for the complex numbers  $\mathbb{C}$ , or a newly defined algebraic system has properties similar to earlier well-studied systems.<sup>23</sup>

Even the arch-deductivist Frege admitted that it is possible to confirm inductively arithmetical theorems by the "countless applications made of them every day". But, as he also pointed out, *if* a deductive proof is available, it is always preferred to "any confirmation by induction".<sup>24</sup> For some famous conjectures, like Fermat's Last Theorem (for  $n \geq 3$ , it is impossible to find non-zero numbers  $x, y, z$  such that  $x^n + y^n = z^n$ ) and Goldbach's Conjecture (every

---

<sup>22</sup> Cf. Niiniluoto (1984), Ch. 8.

<sup>23</sup> For non-demonstrative arguments in set theory, see Maddy (1988).

<sup>24</sup> See Frege (1959), p. 2.

even number except 2 is the sum of two primes), a proof has not been found, and the best evidence for them arises inductively from their known instances.<sup>25</sup>

It is in principle possible to apply systems of inductive logic, designed originally for inferences in empirical science, also in the mathematical domain. If  $g$  is a generalization  $\forall x (Px \rightarrow Qx)$ , and if  $e_n$  describes its instances  $Pa_1 \& Qa_1, \dots, Pa_n \& Qa_n$ , known by proof, then Hintikka's (1966) system defines the inductive probability  $P(g/e_n)$  of  $g$  on  $e_n$ . This probability increases with  $n$ , but decreases with a parameter  $\alpha$ , which expresses the irregularity of the domain of investigation or the caution of the investigator. In mathematical applications, it might be reasonable to choose a small value for  $\alpha$ .

However, it is in order to emphasize that the inductive relation between a general mathematical claim and its instances is a relation between propositions about World 3 entities. Therefore, it need not have anything to do with sense experience. The "quasi-empiricist" account of non-demonstrative reasoning does not give support to a genuinely empiricist philosophy of mathematics.

### 5. *Applications of Mathematics*

It is often thought that the applicability of mathematics to reality is a mystery. The founders of modern mathematical physics were inspired by Platonism: they suggested that God is a geometer who created the physical world according to geometrical ideas. The book of nature is written in a mathematical language, argued Galileo. Another version of the view that mathematical and physical reality are in some sense co-created has been defended by radical "constructivists", like Per Martin-Löf, who think that the world is our construction.<sup>26</sup> On the other hand, the logical empiricists, who thought the propositions of mathematics are devoid of all factual content, claimed that mathematics has only the logically dispensable function of "a theoretical juice extractor" in the establishment of empirical knowledge.<sup>27</sup>

Problems arise also from the tension between the corrigibility of all factual statements and the incorrigibility of mathematical statements: the numerical equation  $2+3=5$  is not disproved by any apparent counter-example.<sup>28</sup>

---

<sup>25</sup> An excellent summary of such evidence is given by Franklin (1987), who also discusses the famous case of the Riemann Hypothesis:

<sup>26</sup> See Martin-Löf (1990). For a critical evaluation of recent sociological forms of "constructivism", see Niiniluoto (1991).

<sup>27</sup> See C.G. Hempel's article in Benacerraf / Putnam (1964), p. 379. Cf. also Field (1980).

<sup>28</sup> See D.A.T. Gasking's article in Benacerraf / Putnam (1964). Cf. also Kim (1981).

Any comprehensive philosophy of mathematics should be able to give an account of the applicability problem. How can this be done, given our thesis that pure mathematics is about World 3?

Let us first observe that quantitative statements may be about World 1 (mathematical physics, mathematical biology) and about World 2 (mathematical psychology). The conditions for applying mathematical concepts to the physical and the mental reality are studied in modern Theories of Measurement.<sup>29</sup> A Representation Theorem establishes a link from World 1 to World 3, or from World 2 to World 3: a factual structure (i.e., a class  $E$  of objects of World 1 or 2, together with some comparative relations between these objects) is mapped into a mathematical structure in World 3 (e.g., the set  $\mathbb{R}$  of real numbers). When the conditions of the Representation Theorem are true, it is guaranteed that the factual structure can be described by using quantitative terms which satisfy ordinary principles of mathematics.

It should be noted that a Representation Theorem asserts the existence of a function  $f: E \rightarrow \mathbb{R}$ , not the existence of  $\mathbb{R}$ . Thus, the Theory of Measurement presupposes that some account (platonist or constructivist) is already given for the reals  $\mathbb{R}$ .

To illustrate with an example how this account solves the applicability problem, let  $E$  be a non-empty set of physical objects,  $\succeq$  a binary relation on  $E$ , and  $\circ$  a binary operation on  $E$ , and define

$$\begin{aligned} a \sim b &\text{ iff } a \succeq b \text{ and } b \succeq a \\ a \succ b &\text{ iff } a \succeq b \text{ and not } b \succeq a \end{aligned}$$

for  $a, b \in E$ . Then the triple  $\langle E, \succeq, \circ \rangle$  is an *extensive system* if

- (i)  $\succeq$  is reflexive, transitive, and connected in  $E$
- (ii)  $a \circ (b \circ c) \sim (a \circ b) \circ c$
- (iii)  $a \succeq b$  iff  $a \circ c \succeq b \circ c$  iff  $c \circ a \succeq c \circ b$
- (iv)  $a \circ b \succ a$
- (v) if  $a \succeq b$ , then for all  $c, d \in E$  there exist a positive  $n \in \mathbb{N}$  such that  $na \circ c \succeq nb \circ d$ , where  $1a = a$  and  $(n+1)a = na \circ a$ .

Then the following theorem can be proved (Krantz *et al.*, 1971):

*Theorem:*  $\langle E, \succeq, \circ \rangle$  is an extensive system iff there exists a positive function  $m: E \rightarrow \mathbb{R}$  such that for all  $a, b \in E$

- 1.  $a \succeq b$  iff  $m(a) \geq m(b)$  and
- 2.  $m(a \circ b) = m(a) + m(b)$ .

---

<sup>29</sup> See Suppes (1967, 1969), Krantz *et al.* (1971).

Function  $m$  is unique up to a similarity transformation: if a positive function  $m': E \rightarrow \mathbb{R}$  satisfies 1 and 2, then  $m' = \alpha m$  for some  $\alpha > 0$ .

Assume now that the elements of  $E$  are physical objects that can be placed in the pans of an equal arm balance. Let  $a \circ b$  mean that both  $a$  and  $b$  are positioned in the same pan,  $a \succeq b$  that the pan with  $a$  is at least as low as the pan with  $b$ , and  $a \sim b$  that the pans with  $a$  and  $b$  are in balance.

Then, at least under certain idealizing assumptions,  $\langle E, \succeq, \circ \rangle$  is an extensive system.<sup>30</sup> The Representation Theorem tells now that there exist a real-valued mass function  $m$  on  $E$  such that  $m(a) \in \mathbb{R}$  is the mass of object  $a \in E$ . Moreover, the embedding of  $E$  into  $\mathbb{R}$  via  $m$  guarantees that the masses of objects obey mathematical laws, e.g.,

$$(7) \quad 2 \text{ kg} + 3 \text{ kg} = 5 \text{ kg}.$$

In a similar way, we can explain why the concept of length satisfies

$$(8) \quad 2 \text{ m} + 3 \text{ m} = 5 \text{ m}.$$

It is essential here that equations (7) and (8) have a physical dimension (such as kilogram, meter). For some other quantities (such as temperature), the corresponding statement is problematic: if in

$$(9) \quad 2^\circ\text{C} + 3^\circ\text{C} = 5^\circ\text{C}$$

we interpret the plus sign  $+$  as indicating that a liquid of temperature  $2^\circ\text{C}$  is placed in the same container as another liquid of temperature  $3^\circ\text{C}$ , sentence (9) turns out to be false.

Representation Theorems also give conditions which show, e.g., when subjective preferences are measurable by cardinal utilities and degrees of belief by personal probabilities.

The applicability of mathematics to the "reality", i.e., Worlds 1 and 2, is thus no mystery. Quantities can be used to represent aspects of reality to the extent that the premises of Representation Theorems are true.

## 6. *Empirical Arithmetic and Geometry*

Arithmetic is not an exception to this general conclusion. When Mill argued that the "fundamental truths" of the "science of number" are "inductive truths" resting on "the evidence of sense", he claimed that propositions concerning numbers "have the remarkable peculiarity that they are propositions

---

<sup>30</sup> Cf. Niiniluoto (1990).

concerning all things whatever, all objects".<sup>31</sup> He illustrated this by a theorem of pebble arithmetic:

- (10) two pebbles and one pebble (put in the same parcel) are equal to three pebbles.

However, principle (10) cannot be generalized to all objects: for example, putting two drops of water and one drop of water in the same cup does not result in three drops.

Hence, a principle like

- (11) 2 objects + 1 object = 3 objects

is true only for objects which *satisfy certain factual conditions*: they do not lose their identity, split or merge with other objects, when collected together as an aggregate. The class of aggregates of such objects, together with the operation  $\circ$  of heaping up, and the comparative relation  $\succeq$  of 'at least as many members as', constitutes an extensive system.<sup>32</sup> The function  $m$ , whose existence is guaranteed by Theorem, is then the cardinality of an aggregate.

Frege is thus perfectly right in pointing out that Mill

"confuses the applications that can be made of an arithmetical proposition, which are often physical and do presuppose observed facts, with the pure mathematical proposition itself".

Nevertheless, pebble arithmetic is *one* of the physical applications where arithmetical equations happen to hold true.<sup>33</sup>

Besides pebbles, *empirical arithmetic* may concern the operation of physical devices, such as pocket calculators or computers. Suppose my calculator informs that  $789 \times 456 = 359784$ . As I have a high degree of confidence in the regular behaviour of my calculator, even this *one* instance leads me to believe that it will always (under normal conditions) give the same answer. Hence, I know that

- (12) According to my pocket calculator,  $789 \times 456 = 359784$ .

This is a physicalist statement about the empirically observable behaviour of a physical object. However, at the same time (12) gives inductive confirmation to the pure arithmetic equation ' $789 \times 456 = 359784$ ', if we have good reasons

<sup>31</sup> See Mill (1906), pp. 167–168.

<sup>32</sup> It has been well-known since Cantor and Frege that the relation 'has at least as many members as' can be defined without counting cardinal numbers.

<sup>33</sup> See Frege (1959), p. 13. For a recent discussion of pebble arithmetic, see Bigelow (1988).

for believing that the physical states and operations of the calculator constitute a model of arithmetic.<sup>34</sup> Similarly, verification of mathematical statements by a computer gives empirical confirmation to their truth.<sup>35</sup> This kind of inductive evidence applies primarily to singular mathematical statements. It satisfies many familiar principles from other domains: for example, two independent witnesses corroborate each other's testimonies.

One important qualification – for which Mill deserves credit – has to be made here. He points out that the "necessity" of the conclusions of geometry "consists in reality only in this, that they correctly follow from the suppositions from which they are deduced". But those suppositions "are so far from being necessary that they are not even true; they purposely depart, more or less widely, from the truth". Thus, deductive sciences (under Mill's empirical interpretation) are inductive and hypothetical: they are based on axioms that "are, or ought to be, approximations to the truth". This does not mean that these axioms are suppositions "not proved to be true, but surmised to be so", but rather they are "known not to be literally true, which as much of them as is true is not hypothetical, but certain".<sup>36</sup>

Mill is right in thinking that the application of mathematics to reality involves simplifying and idealizing assumptions, so that the interpreted mathematical statements are at best known to be approximately true.<sup>37</sup> His intuitive ideas can be made precise by using tools of modern logic.<sup>38</sup> Mill also makes sensible remarks about the confirmation or "proof by approximation" of the axiom that "two straight lines cannot inclose a space". What he does not notice is that, within sufficiently small regions, alternatives to the Parallel Axiom may be approximately true as well. The kind of empirical evidence that Mill appeals to is not sufficient to decide between Euclidean and non-Euclidean physical geometry.

---

34 Jon Rinden (1980) has argued that transformational generative grammar is a "non-empirical discipline like logic, pure mathematics, or formal analytical philosophy", since it is analogous to the search and evaluation of the axioms of arithmetic. However, his description of arithmetic equations corresponds to empirical arithmetic – so that his analogy between linguistics and mathematics in fact supports the opposite of his own conclusion (cf. Niiniluoto, 1981).

35 Cf. Tymoczko's account of the 4CT.

36 See Mill (1906), p. 149.

37 Similarly, Kitcher (1983) takes his "Mill Arithmetic" to be "an idealizing theory" about "ideal operations performed by ideal agents".

38 For concepts of truthlikeness and approximate truth, applicable also to idealized theories, see Niiniluoto (1984, 1986, 1987, 1991).

## 7. *Mathematical and Empirical Theories*

The observations above allow us to make some general remarks about the relation between mathematical and empirical theories.

Patrick Suppes (1967) coined the slogan that to axiomatize a theory is to "define a set-theoretical predicate". His paradigm case was the algebraic theory of groups: a pair  $\langle G, o \rangle$ , where  $G$  is a non-empty set and  $o$  is a binary operation in  $G$ , is said to be a *group* if and only if  $o$  is associative, there is an identity element  $u$  in  $G$ , and every element  $a$  in  $G$  has an inverse  $a^{-1}$ .

Joseph Sneed and Wolfgang Stegmüller extended Suppes' account by introducing the concept of *intended model*; Stegmüller (1979) also suggested that the new "structuralist" view of empirical theories is an analogue of Bourbaki's treatment of mathematical theories. However, in addition to Wolfgang Balzer's work on empirical geometry and empirical arithmetic, the structuralist school has not analysed theories of pure mathematics.<sup>39</sup>

Let  $A$  be the set of axioms of a mathematical theory, and let  $I$  be the class of the intended models of this theory. Then (simplifying the structuralist account) a *theory* can be defined as the pair  $T = \langle A, I \rangle$ ; the *claim* of theory  $T$  is that each element  $M$  of  $I$  satisfies the axioms  $A$ , i.e.,  $M \models A$  for all  $M \in I$ .

Pure mathematics includes theories of two types. The first type is exemplified by Group Theory: its intended models include all structures in World 3 which are groups (e.g., integers with addition). Its claim is thus a trivial analytic truth: every group is a group. Progress in the study of such a theory means the deduction of new consequences from the axioms and, thereby, the gain of new insight of the variety and classification of different kinds of groups.<sup>40</sup>

The second type is exemplified by Arithmetic. It has a *standard model*  $\mathbb{N}$  which is unique up to isomorphism. The claim of the theory  $\langle \text{Peano Arithmetic, structures isomorphic to } \mathbb{N} \rangle$  is again an analytic truth. Progress for such a theory means that we derive more and more informative truths about the intended model  $\mathbb{N}$ .

We have seen in Sections 5 and 6 that it may be possible to interpret, via Representation Theorems, mathematical theories also in Worlds 1 and 2. The statement that a physical or mental structure satisfies quantitative axioms is non-analytic, based upon facts about World 1 or 2.

Hence, it is possible to construe empirical arithmetic as a theory  $\langle \text{PA}, I \rangle$ , where the intended applications in  $I$  belong to World 1 only. This is what

39 Cf. Balzer *et al.* (1987). Hellman's (1989) "structural" approach does not mention the Sneed–Stegmüller "structuralism" at all.

40 For the concept of progress in mathematics, see Niiniluoto (1984), Ch. 8.



Balzer (1979) does in choosing as the members of  $I$  "those finite sets of concrete objects which can be counted by human beings".<sup>41</sup> It is not clear, however, that this constitution of numbers dispenses with abstract entities, since a *set* of concrete objects is not a concrete object. Further, it is not plausible that such a physicalist treatment captures the whole content of arithmetic, since in World 1 there are not arbitrarily large numbers of concrete objects. To express the full content of arithmetic, reference to World 1 only is not sufficient.

Indeed, one and the same theory may make a pure mathematical claim about World 3 *and* empirical claims about Worlds 1 and 2. The class  $I$  of intended models of such a theory is then divided into three disjoint subclasses

$$I = I_1 \cup I_2 \cup I_3,$$

so that structures in  $I_i$  belong to World <sub>$i$</sub>  ( $i = 1, 2, 3$ ). The claim of such a theory is then likewise divided into a factual and an analytic part.

It is also possible that a quantitative theory has "mixed" interpretations. For example, a theory of psychophysics asserts something about the interrelations between elements of World 1 and World 2. A theory in the cultural or social sciences may have a model which exhibits interrelations between elements from each of the three worlds.

### Bibliography

- Balzer, W., "On the Status of Arithmetic", *Erkenntnis* 14 (1979), 57–85.  
 Balzer, W. / Moulines, C. U., / Sneed, J. D., *An Architectonic for Science: The Structuralist Approach*, D. Reidel, Dordrecht, 1987.  
 Benacerraf, P./ Putnam, H. (eds.), *Philosophy of Mathematics*, Prentice-Hall, Englewood Cliffs, N. J., 1964. 2nd ed. Cambridge UP, Cambridge, 1983.  
 Bigelow, J., *The Reality of Numbers: A Physicalist's Philosophy of Mathematics*, Oxford UP, Oxford, 1988.  
 Dummett, M., "Wittgenstein's Philosophy of Mathematics". In: Benacerraf and Putnam (1964), 491–509.  
 Field, H., *Science without Numbers*, Princeton UP, Princeton, 1980.  
 Franklin, J., "Non-deductive Logic in Mathematics", *The British Journal for the Philosophy of Science* 38 (1987), 1–18.  
 Frege, G., *The Foundations of Arithmetic*, Blackwell, Oxford, 1959.  
 Gillies, D. A., "Intuitionism versus Platonism: A 20th Century Controversy Concerning the Nature of Numbers". In: F. Gil (ed.), *Scientific and Philosophical Controversies*, Fragmentos, Lisbon, 1990, 299–314.  
 Hellman, G., *Mathematics without Numbers: Towards a Modal-Structural Interpretation*, Oxford UP, Oxford, 1989.

---

<sup>41</sup> Cf. also Kitcher (1983) and Bigelow (1988).

- Hersh, R., "Some Proposals for Reviving the Philosophy of Mathematics". In: Tymoczko (1986), 9-28.
- Hintikka, J., "A Two-Dimensional Continuum of Inductive Methods". In: J. Hintikka and P. Suppes (eds.), *Aspects of Inductive Logic*, North-Holland, Amsterdam, 1966, 113-132.
- Irvine, A. D. (ed.), *Physicalism in Mathematics*, Kluwer Academic Publishers, Dordrecht, 1990.
- Kim, J., "The Role of Perception in *a priori* Knowledge: Some Remarks", *Philosophical Studies* 40 (1981), 339-354.
- Kitcher, P., *The Nature of Mathematical Knowledge*, Oxford UP, Oxford, 1983.
- Krantz, D. / Luce, R. / Suppes, P. / Tversky, A., *Foundations of Measurement*, vol. 1, Academic Press, New York, 1971.
- Lakatos, I., *Proofs and Refutations*, Cambridge UP, Cambridge, 1976.
- Maddy, P., "Believing the Axioms", *Journal of Symbolic Logic* 53 (1988), 481-511, 736-764.
- Martin-Löf, P., "A Path from Logic to Metaphysics", a paper presented at a congress in Viareggio, January 1990.
- Mill, J. S., *A System of Logic*, Longmans, Green, and Co., London, 8th ed., 1906.
- Niiniluoto, I., "Language, Norms, and Truth". In: I. Pörn (ed.), *Essays in Philosophical Analysis*, Acta Philosophica Fennica 32 (1981), 168-189.
- Niiniluoto, I., *Is Science Progressive?*, D. Reidel, Dordrecht, 1984.
- Niiniluoto, I., *Truthlikeness*, D. Reidel, Dordrecht, 1987.
- Niiniluoto, I., "Theories, Approximations, and Idealizations". In: R.B. Marcus, G.J.W. Dorn, and P. Weingartner (eds.), *Logic, Methodology and Philosophy of Science VII*, North-Holland, Amsterdam, 1986, 255-289. (Reprinted and extended in J. Brzezinski et al. (eds.), *Idealization I: General Problems*, Rodopi, Amsterdam, 1990, 9-57).
- Niiniluoto, I., "Realism, Relativism, and Constructivism", *Synthese* 89 (1991), 135-162.
- Niiniluoto, I. / Tuomela, R., *Theoretical Concepts and Hypothetico-Inductive Inference*, D. Reidel, Dordrecht, 1973.
- O'Hear, A., *Karl Popper*, Routledge & Kegan Paul, London, 1980.
- Parsons, Ch., "The Structuralist View of Mathematical Objects", *Synthese* 84 (1990), 303-346.
- Polya, G., *How to Solve It: A New Aspect of Mathematical Method*, Princeton UP, Princeton, 1945.
- Polya, G., *Induction and Analogy in Mathematics*, Princeton UP, Princeton, 1954.
- Popper, K., *Objective Knowledge*, Oxford UP, Oxford, 1972.
- Popper, K. / Eccles, J. C., *The Self and Its Brain*, Springer-International, Berlin, 1977.
- Putnam, H., "What is Mathematical Truth?". In: *Mathematics, Matter and Method*, Cambridge UP, Cambridge, 1975, 60-78. (Reprinted in Tymoczko, 1986, 49-65).
- Reichenbach, H., *The Philosophy of Space & Time*, Dover, New York, 1957.
- Resnik, M., "Mathematics as a Science of Patterns: Ontology and Reference", *Noûs* 15 (1981), 529-550.

- Ringen, J., "Linguistic Facts: A Study of the Empirical Scientific Status of Transformational Generative Grammars". In: T.A. Perry (ed.), *Evidence and Argumentation in Linguistics*, Walter de Gruyter, Berlin-New York, 1980, 97-132.
- Simons, P. M., "Three Essays in Formal Ontology". In: B. Smith (ed.), *Parts and Moments*, Philosophia Verlag, München, 1982, 111-260.
- Stegmüller, W., *The Structuralist View of Theories: A Possible Analogue of the Bourbaki Programme in Physical Science*, Springer-Verlag, Berlin-Heidelberg-New York, 1979.
- Stenius, E., *Critical Essays*, Acta Philosophica Fennica 25, North-Holland, Amsterdam, 1972.
- Suppes, P., *Set-Theoretical Structures in Science*, Institute for Mathematical Studies in the Social Sciences, Stanford University, Stanford, 1967.
- Suppes, P., *Studies in the Methodology and Foundations of Science*, D. Reidel, Dordrecht, 1969.
- Tait, W.W., "Truth and Proof: The Platonism of Mathematics", *Synthese* 69 (1986), 341-370.
- Troelstra, A. S., *Principles of Intuitionism*, Springer-Verlag, Berlin-Heidelberg-New York, 1969.
- Tuomela, R., *Theoretical Concepts*, Springer-Verlag, Wien-New York, 1973.
- Tymoczko, T. (ed.), *New Directions in the Philosophy of Mathematics*, Birkhäuser, Boston-Basel-Stuttgart, 1986.

## Tacit Knowledge in Mathematical Theory

HERBERT BREGER (Hannover)

Let me begin with three problems that will establish the scope of my topic. First there is the problem of logicism. Many philosophers, especially analytical philosophers, follow one or other version of logicism, according to which investigation of logical structure makes the essential contribution to understanding that which is mathematics. According to Bertrand Russell mathematics is just a branch of logic. Among mathematicians neither Russell's strong version nor the weak version common today has met approval. Mathematicians simply know that mathematics is something essentially different from logic. The logicistic philosophers have on the whole, however, an easy time since mathematicians cannot substantiate their conviction or do not wish to substantiate it. If a discussion takes place at all between a logicistic philosopher and a working mathematician, then a bad compromise is usually the outcome: the philosopher makes clear that he is not concerned with the creative process in the spirit of the mathematician but rather with the analysis of proofs and the reconstruction of mathematics on the basis of a logicistic concept of proof. The mathematician is relieved to hear this; somehow his reservations are taken account of and he does not need to involve himself in an uncomfortable discussion. The compromise does not, however, go to the heart of the matter; the essential point is neither the creative process nor the trivial fact that proofs have a logical structure but rather the structure and construction of mathematical theory as well as the direction of its development. Granted, a poem results from the application of orthography and one may therefore consider poetry to be a branch of orthography, but it is not very convincing to characterise a complex system by its most uninteresting aspect. That which interests me here is, however, not *only* the inappropriateness of the different varieties of logicism but also the difficulties mathematicians have in expressing their aversion to logicism in the form of arguments.

The second problem with which I wish to mark out the scope of my topic is the philosophical problem of the history of mathematics. Anyone who

concerns himself profoundly with the mathematics of a past century will have again and again a confusing experience: the texts are familiar and at the same time very remote; they appear to be referring to the same topic as a modern textbook and yet to deal with different matters. One can rid oneself of one's own confusion by training oneself to automatically translate that which one reads into the sphere of the ideas and concepts of modern mathematics. But, just as in the case of a literary text, something of the spirit of the original language is lost in translation with, at the same time, new resonant associative contents arising, so also, in the translation of historical mathematical texts into the terminology of today's mathematics, the specific character of the historical text is lost. The content of an historical mathematical text appears to consist not only in that which is explicitly stated in it but also in an implicit background knowledge, the atmosphere so to speak of the mathematics of a past epoch. Sometimes translation into the conceptual and thought spheres of modern mathematics gives the impression that important mathematicians of an earlier century did not do what they should have done. Without doubt their results are correct, but the path along which they have obtained some of their results would, if followed, lead a student nowadays to certain failure in examination. Are we really to suppose that Fermat, Leibniz and Poncelet were somewhat confused in their mathematical thought, or is it perhaps the case that we have a simplified vision of mathematical progress?

The third problem that I would like to refer to at the outset concerns the abilities of a computer in the field of mathematics. If we feed a computer with the axioms of topology (including the axioms for logic and set theory), it can in principal derive topological theorems. But the computer is in fact in no way in a position to write a textbook of topology. It possesses no criteria for differentiating between interesting and trivial statements; a fundamental theorem means nothing more to the computer than some correct line or other in a proof. How is this incapability of the computer possible? Doesn't everything follow from the axioms? Now mathematics consists not just in logical deductions but, above all, in the ability to differentiate *from within the area of correct statements* the elegant, profound and essential statements from the uninteresting statements. The computer cannot make such value judgements because there are no universally valid criteria and no explicit definitions for "important" or "profound" and so on. The fact that specialists are in agreement as regards decisions in relation to this is of the greatest importance here; these are not arbitrary, subjective decisions but rather real objective knowledge – granted knowledge that cannot be explicitly expressed in criteria or definitions. This "tacit knowledge" or "know-how" is the subject of the following

considerations. Knowledge of formal deductions could, in contrast, be designated as explicit knowledge or "know-that". This terminology has been introduced<sup>1</sup> by Michael Polanyi and the Dreyfus brothers and is used in the philosophical debate about artificial intelligence and the performance capability of computers. This terminology is also useful in the philosophy of mathematics since there is also in mathematics a knowledge that cannot be made explicit in rules and definitions and accordingly is non-programmable. Tacit knowledge is that knowledge that differentiates between the specialist in one particular area of mathematics and a student who can only understand the individual steps of the logical deductions. One can define a specialist in one branch of mathematics by the fact that he is able to correctly apply words like "elegant", "simple", "natural", "profound" in this branch of mathematics, although there are no universally applicable rules for the use of such words. For the very reason that these words are undefined, they are suitable for evoking that tacit knowledge in conversation between specialists that provides the understanding of the formal theory. Of course, tacit knowledge of a theory also encompasses the ability to solve easy problems or those of intermediate difficulty. For mathematical research tacit knowledge is of course a prerequisite; I would like, however, to confine myself to the established region of mathematics since I do not want to burden my considerations with unclear and equivocal concepts such as "creative process" or "creative intuition". In the psychology of invention<sup>2</sup> one does not have at one's disposal the criterium of the unanimity of specialists, with the result that it would be considerably more difficult to find a demarcation from purely subjective thoughts and conceptions.

In teaching the difference between know-how and know-that becomes clear. Whereas know-that can be explicitly written on the blackboard and the student only needs to write it down or to employ his memory, know-how cannot be written on the blackboard. It can only be taught by doing or demonstration and the student must obtain an understanding of the matter by his own activity. Riding a bicycle is a good example: one cannot learn it even from the most detailed verbal instruction but only by trying it oneself (although undoubtedly verbal instructions can be useful). Mathematicians therefore know something that they are not able to communicate. This might lead one to fear that in this

---

<sup>1</sup> Michael Polanyi, *Personal Knowledge*, London, 1962, especially p. 124–131, p. 184–193; Hubert Dreyfus/Stuart Dreyfus, *Mind over Machine*, New York, 1986. Cf. also Breger "Know-how in der Mathematik", in: Detlef Spalt (ed.): *Rechnen mit dem Unendlichen*, Basel, 1990.

<sup>2</sup> On this subject cf. Jacques Hadamard, *An Essay on the Psychology of Invention in the Mathematical Field*, Princeton 1945.

way mathematics becomes similar to mysticism: only through lengthy exercise, during which that which is most important cannot be said, does one become initiated and, just as the mystic speaks in metaphors, the mathematician uses words like "elegant", "beautiful" and "natural". But mathematics is without doubt a pretty rational undertaking and, in the last analysis, no more mystical than riding a bicycle.

At this point two objections are to be expected. Firstly, a logician will object that no differentiation is made between mathematical theory and the metalevel; tacit knowledge concerns apparently speaking *about* mathematical theory. The objection shows, however, only the difficulty of understanding mathematics from a logicistic standpoint; mathematical understanding reveals itself only on the metalevel. Differentiation between formal theory and the metalevel has a good purpose but "mathematics" may under no circumstances be equated with formal theory that really exists on paper. All decisions concerning structure, construction and further development of a mathematical theory are taken on the metalevel. The second objection consists in pointing out that there is little point in talking about that which cannot be communicated. One must be silent, according to Wittgenstein, about that about which one cannot speak. But mathematical progress consists partly in the fact that parts of the tacit knowledge of an epoch become, in the course of an historical process, more and more familiar and self-evident and then can in the end be made explicit and formalised. In such cases vague but fruitful and familiar ideas and abilities are admitted from the metalevel into the formal theory; the metalevel thus contains the air needed by the formal theory for breathing, living and development. If one excludes the metalevel, then one is indulging in the anatomy of mathematics and thus the dissection of the corpse. In the following examples from the history of mathematics there exists of course the methodical problem of how to show that a mathematician of an earlier epoch may have known something that he did not make explicit. I hope, however, to be able to make this sufficiently plausible.

I would like to differentiate between a number of types of tacit knowledge; even where the borders between them are not distinct they may perhaps, in their totality, contribute to a better understanding of mathematical progress. The first type is the insight and the understanding of a theory, thus of that knowledge that a mathematician has in advance of the computer programmed with axioms, in particular for example knowledge of the different relevance of correct statements of theory. Directly connected with this first type of tacit knowledge is the second type, the know-how for axiomatisation. Only when a theory, that in the first place is not axiomatically constructed, has developed to the point where a comprehensive and in a certain sense closed understanding of

its internal structure becomes possible, can it be axiomatised. The development of algebraic topology offers an instructive example. The *Lehrbuch der Topologie* (Textbook of Topology) by Seifert / Threlfall of 1934 shows the state of the theory before axiomatisation. The homology groups are constructed and worked out in steps; the transition from the topological space to the homology groups is constructed. The book contains numerous illustrations that appeal to the geometric power of imagination of the reader or attempt to bring it out in the mind of the student. He who is familiar with the theory possesses a highly developed geometrical power of imagination as well as a feeling and faculty of judgement for the construction of the theory, the relevance of the individual propositions as well as the central ideas and methods of proof. Just this tacit knowledge of the specialists is the starting point for the axiomatisation undertaken by Eilenberg/Steenrod in 1952 in *Foundations of Algebraic Topology*. Different definitions for homologies had established themselves: the singular homology groups of Veblen, Alexander and Lefschetz, the relative homology groups, the Vietoris homology groups, the Čech homology groups etc. "In spite of this confusion, a picture has gradually evolved of what is and should be a homology theory. Heretofore this has been an imprecise picture which the expert could use in his thinking but not in his exposition".<sup>3</sup> Through the axiomatisation, that is the transition to a higher level of abstraction, a precise picture emerges from the tacit knowledge. The impression that an "imprecise" picture had previously existed arises only when the higher stage of abstraction has already been attained. Each and every one of the axioms formulated by Eilenberg and Steenrod is a theorem of classical homology theory, but in most cases it is not clear who first stated and proved them.<sup>4</sup> Some of the axioms are too trivial as that one should have thought of expressly stating them in the older theory. Another axiom, the excision property, had been implicit in Lefschetz's construction of the relative groups. Similarly the group homomorphism, that is induced by a continuous map between topological spaces, and the boundary operator had been used for a long time without formal recognition. The formal recognition of a concept, already used for a long time, implies *seeing* the theory in a new way, or, in other words, a transformation in the value judgements connected with the theory, as for example in the decisions about what is essential. Above all it is the value judgements that determine the atmosphere of a theory. The book of Eilenberg and Steenrod does not contain a single illustration that is directed at the

---

<sup>3</sup> Samuel Eilenberg / Norman Steenrod, *Foundations of Algebraic Topology*, Princeton, 1952, p. VIII.

<sup>4</sup> Ibid. p. 47.



development of the geometrical power of conception of the reader; all illustrations are diagrams with commutativity relations and exact sequences. In the old theory that which provided geometrical insight was considered "natural" and "beautiful". In the new theory beauty consists more in the simplicity of the algebraic machinery. Whereas, on the one hand, trivial statements of the older theory are expressly formulated as axioms in the new theory, on the other hand normal proofs of the older theory appear clumsy in the new theory. Interesting theorems of the older theory now appear in the new theory as application examples and exercises in the appendix. In a certain sense the older theory is contained in the newer theory, but nevertheless it is true in another sense that the older theory communicated more geometrical knowledge. If the acquisition of *this* knowledge were the real objective of topology, then the new theory would be an unnatural and complicated detour.

In as much as the homology theory converts geometrical circumstances into algebraic, it has similarity with analytical geometry. I would therefore like to consider the *Géométrie* of René Descartes, of the year 1637, as the next example.<sup>5</sup> What knowledge did the specialist in the area of Euclid's geometry at the beginning of the 17th century have? He had at his disposal not only a certain know – that of the properties of straight-lines, circles, secants, similar triangles, etc., but also a specific proficiency and cunning in the organisation of this knowledge for the solution of given problems by means of construction with ruler and compass. Many such problems had been posed and solved by Euclid but reprints of Euclid editions in the early modern period contain also, as a rule, new problems and their solutions. Moreover, the conviction had become more or less widespread that certain problems (as, for example, the trisection of an angle) could not be solved with ruler and compass. Descartes now showed how construction problems can be solved without being in possession of any special proficiency and cunning. One makes a drawing and marks the given and required magnitudes with letters. Then one establishes algebraic relations between these magnitudes, in the course of which the theorem of Pythagoras, theorems concerning similar triangles and the like are useful. One combines these relationships in the form of an equation for the unknown and solves this equation. This algebraic solution can now be interpreted at once as a construction rule for the original geometrical problem, because the sum, difference, product and quotient of two lengths, as well as the square root of a given length, can be constructed with ruler and compass. Descartes therefore provides a general procedure by means

---

<sup>5</sup> Cf. Henk Bos: "The Structure of Descartes' *Géométrie*". In: *Descartes: il Metodo e i Saggi*, vol. 2, Rome, 1990, pp. 349–369.

of which the specific know-how for the solution of problems with ruler and compass becomes superfluous. Strictly speaking, his procedure is not completely formalised; in the first step one still requires a little elementary know-how; one must for example be able to draw a suitable auxiliary line or the like.

A good part of mathematical progress in the early modern period comes about through the formalisation of know-how for problem solving. Thus the know-how for the solution of problems in number theory, as presented by Diophantus in his "Arithmetic", is formalised through Viète's introduction of calculating using letters. A typical problem of Diophantus is the following: find three numbers such that one obtains given numbers on multiplying the sum of any two by the third.<sup>6</sup> Diophantus knows how one solves this problem but he cannot completely communicate this knowledge. He is in possession of a special sign for the first unknown but not for the second and third unknown and not for given numbers. As a consequence he can demonstrate his ability to solve the general problem only exemplarily with a particular numerical example. For the second and third unknowns he assumes numbers arbitrarily and then calculates until he can see how this arbitrary assumption has to be corrected. This correction can of course only then be carried out if one has in one's mind how the numbers attained have come about through addition and multiplication from the given numbers. In other words, Diophantus allows us to look on at how he cooks the dish, and François Viète writes the cookbook.

Fermat's procedure for the solution of extreme-value problems is an instructive example of misunderstandings in the history of mathematics. On the one hand Fermat has been seen as the discoverer of the differential calculus,<sup>7</sup> on the other hand he has continually faced accusation that he is unable to prove the correctness of his procedure, said to be based on the equating of expressions that at all events could only be approximately equal, in short, that his procedure is a mystery. In order to understand Fermat, one must take notice of the fact that he presents his procedure using examples and answers objections that his procedure is dependent on chance with the formulation of a new problem for which his procedure is likewise successful.<sup>8</sup> Fermat apparently does not at all have the intention of providing a deductive, proving theory of extreme values; he shows rather how one can find the extreme values of particular curves. Strictly speaking one still has to prove, in

---

6 Diophant, *Opera omnia*, ed. by P. Tannéry, vol. 1, Leipzig, 1893, p. 216–221.

7 Cf. Margaret Baron, *The Origins of the Infinitesimal Calculus*, Oxford–London etc., 1969, p. 167.

8 Pierre de Fermat, *Varia opera mathematica*, Toulouse, 1679, p. 153.

accordance with the traditional scheme of analysis and synthesis, that the value obtained is really an extreme value. But this proof is not very difficult for the specialist; Fermat thus considers it not worth the trouble of writing it down. Thus the objections against Fermat and the apparent mysticism are founded on the confusing of two different levels of abstraction. Fermat is interested in individual problems of curves and, through his intimate familiarity with these problems, has gained a certain tacit knowledge, namely the correct conviction that it is possible to prove, for each individual case in which the procedure is applicable at all, that the value provided by the procedure is really an extreme value. Fermat's conviction is a matter of the metalevel and has nothing to do with the logical correctness of his mathematics since proof was actually possible in each and every individual case. Today we have available an explicit conceptuality and a formal theory on an abstract level on which we can formulate, and prove once and for all, theorems of extreme values and a concept such as "derivation". We no longer need the intimate familiarity with problems of curves with the result, however, that we also no longer "see" what Fermat "saw". Furthermore, the accusation often made that Leibniz's infinitesimal calculus rests on an unsound foundation is inadmissible for similar reasons; I will not go into this now since I have dealt with the matter elsewhere.<sup>9</sup>

For the present-day mathematician it is, in the light of these problems, astonishing how stubbornly the lower level of abstraction was maintained. This apparently is because the objects of investigation were considered as directly given or as "natural". Concepts like curve, tangent or area under a curve were not axiomatically introduced; thus problems attain a systematic priority and the transition to more abstract modes of conception takes place only hesitatingly. A surprisingly late example of this traditional way of thinking is the Erlangen Programme of Felix Klein. The concept of the group of transformations is designated in the Erlangen Programme as its "most essential concept"<sup>10</sup> and yet this concept is not, or rather incorrectly, defined. Neither the existence of an identity nor of an inverse nor the associativity are mentioned. Decades later, in his lectures on the development of mathematics in the 19th century, Klein does refer to the precise definition of a group but

---

<sup>9</sup> H. Breger, "Le continu chez Leibniz", to be published in the proceedings of the conference "Le continu mathématique" (Cerisy-la-Salle, 11.-21.9.1990) edited by H. Sinaceur and J. M. Salanskis.

<sup>10</sup> Felix Klein, "Vergleichende Betrachtungen über neuere geometrische Forschungen". In: *Gesammelte mathematische Abhandlungen*, vol. 1, Berlin, 1921, p. 462.

with an unmistakable warning against too extensive abstraction.<sup>11</sup> Apparently for Klein it was the geometry in the Erlangen Programme that was the object of investigation; the group concept could continue to be, to a certain extent, vague and self-evident, since it was simply a means of establishing order in a given area of investigation. A precise definition did indeed then become necessary when the groups themselves, be it as permutation groups or as groups of geometrical transformations, were made the object of investigation. An accusation of unclear thought or inadequate mathematical rigour would appear to be just as much out of place in the case of Klein as in the case of Fermat or that of Leibniz.

In pure mathematics of the 20th century things look different. The objects of investigation (at least with the exception of elementary number theory) are introduced axiomatically; within formal theory intuitive convictions only play a role at a few and sharply defined points: Gödel's theorem, the thesis of Church and the form principle in the geometry of Paul Lorenzen should be mentioned. Whenever a class of problems arises an attempt is made in today's pure mathematics to go over at once to a higher level of abstraction; just this rapid transition appears to be the characterising style of thought of pure mathematics in the 20th century. Accordingly the importance of the formalisation of know-how for problem solving has waned as the driving force of mathematical progress; today the driving force appears rather to be the formalisation of know-how for finding the right definition, the right construction or the right generalisation.<sup>12</sup> This know-how appears to be a fourth type of tacit knowledge.

Prominent examples of this fourth type are provided in great number by the theory of categories. With the help of the concept of the adjoint functor, the right definition or right generalisation can be given; I will confine myself to an example from topology.<sup>13</sup> In the category of topological spaces and continuous maps an exponential law for function spaces is valid if one of the two factors of the product is a locally compact Hausdorff space. But this precondition is too strong to be convenient. Various attempts have been made to attain a general validity of an exponential law by providing the product space or the function space with a topology other than the ordinary one. The

---

11 Felix Klein, *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*, vol. 1, Berlin 1926, p. 335–336.

12 Cf. the examples in Andreas Dress, "Ein Brief". In: Michael Otte (ed.): *Mathematiker über die Mathematik*, Berlin–Heidelberg–New York, 1974, p. 161–179.

13 Cf. Saunders Mac Lane, *Categories for the Working Mathematician*, New York–Heidelberg–Berlin, 1971, p. 181–184.

best solution was found by Kelley, in 1955, with the definition of compactly generated Hausdorff spaces. With the help of the category theory it is easy to see that this is in a certain sense the right definition. Furthermore category theory also shows what has to be done if one wishes to generalise Kelley's definition and to drop the prerequisite "Hausdorff". Kelley found his definition a few years before the formulation of the concept of adjoint functor on the lower level of abstraction by means of topological insight into the problem of this special case. It would apparently be misleading if one wanted to say in the case of Kelley, analogous to the previously mentioned criticism of Fermat, that he found his definition in a mysterious and unclear manner, although it is in fact the case that he could not prove that his definition is the best.

The concepts "category" and "functor" have been introduced in order to be able to define "naturality"; the first papers of Eilenberg and Mac Lane had the title "Natural Isomorphism in Group Theory" and "General Theory of Natural Equivalence".<sup>14</sup> Eilenberg and Mac Lane illustrated their idea with the example of isomorphism between a vector space and its dual space: The isomorphism is not natural, since it depends on the choice of a special base, but the isomorphism between a vector space and the dual of the dual space is natural. Accordingly category theory is the successful attempt to formalise, at least partly, one of the undefined words on the metalevel by which allusion is made to tacit knowledge. This definition of "natural" offers a kind of guide for the formation of concepts in the abstract parts of mathematics, like that offered by the intuitive naturality in the pure mathematics of the 19th century and even today in applied mathematics.<sup>15</sup> The experiencing of missing intuitive naturality becomes clear in 1914 in Hausdorff's book on general topology, that at that time was beginning to separate itself from set theory, when Hausdorff writes in the foreword that he is dealing with a "territory where plainly nothing is self-evident, the correct often paradox and the plausible false".<sup>16</sup> The founders of the theory of categories certainly echoed a mood found among mathematicians around 1950 when they with self-irony called their own idea

---

<sup>14</sup> *Proceedings of the National Academy of Science of the USA* 28, 1942, p. 537–543, resp. *Transactions of the American Mathematical Society*, 58, 1945, p. 231–294. Cf. also Mac Lane, "Categorical Algebra", *Bulletin of the American Mathematical Society* 71, 1965, p. 48.

<sup>15</sup> Cf. David Ruelle, "Is our Mathematics Natural? The Case of Equilibrium Statistical Mechanics". *Bulletin of the American Mathematical Society* 19, 1988, p. 259–268.

<sup>16</sup> Felix Hausdorff, *Grundzüge der Mengenlehre*, Leipzig, 1914, p. V ("Gebiet, wo schlechthin nichts selbstverständlich und das Richtige häufig paradox, das Plausible falsch ist"), cf. also p. 211, p. 369, p. 469–472.

"abstract nonsense". Meanwhile, the new feeling for naturality, now founded on a formal definition, is long established; in a textbook of 1966 we find brief and to the point the statement: "Particular emphasis has been placed on naturality, and the book might well have been titled *Functorial Topology*".<sup>17</sup> In addition to this formally defined concept of naturality, there continues to exist the undefined use of "natural" on the metalevel, at least in other parts of mathematics.

In conclusion I would like to mention a last type, namely the tacit knowledge of the trivial. A mathematical proof is not that what a logician understands under proof. Mathematics would become cumbersome and extremely tedious if all proofs were to be written out in full.<sup>18</sup> Rather it is the case that routine arguments, and all which is obvious to the presumed reader, are simply omitted. That which is considered trivial by the specialists in a particular area may be as good as unintelligible to a specialist in another area of mathematics. That which is trivial can through the further development of mathematics cease to be trivial as has already been mentioned in the case of the axiomatisation of algebraic topology and that of Felix Klein's definition of the group of transformations. On the other hand a clever trick, that is successful in a particular case, can become in the course of the further development a routine method that is applied in many cases and scarcely appears worth mentioning. Or a theorem that is difficult to prove can, as a result of the development of a general theory, become a trivial consequence of this theory. Decisive is now that there is no definition of the trivial by means of the formal theory. Furthermore there are no criteria by which for example a computer could decide whether an existing gap, from a logical point of view, in a proof in a journal article could be filled by a trivial consideration or whether a real mistake in the proof exists. Certainly, for each branch of mathematics, a list of the most frequently occurring trivialities could be made, but such a pragmatically prepared list will not be complete. As a complete list one can consider the totality of all previously proved theorems of the theory, but such a list is of very little value in writing a computer programme since it says nothing as to the combination or as to the aspects of the case in question to which the theorems proved previously must be applied in order to fill the gap in the proof. At all events there does not yet exist such a programme. The mathematician does not have at his disposal a general theory of that which is

---

<sup>17</sup> Edwin H. Spanier, *Algebraic Topology*, New York–San Francisco etc., 1966, p. VII.

<sup>18</sup> Philip Davis / Reuben Hersh, *Descartes' Dream*, San Diego–Boston–New York, 1986, p. 66, p. 68–69, p. 73.

omissible, but simply drops in certain individual cases certain steps that are necessary for the proof. This is a matter of the decisions of the mathematician on the metalevel that intervene with the course of the mathematical proof. We are all convinced that the gaps in the proofs could be filled but they are, at all events, not filled and the general conviction rests on tacit knowledge. This certainly does not mean that all proofs stand on uncertain foundations, for, in each individual case, a possible sceptic can fill the gaps or establish the existence of a mistake. Likewise we cannot formalise our ability to ride a bicycle in rules and still only fall off in exceptional cases, and what is even more comforting is that, when we do fall off, we are in the aftermath able to give an explanation for it. In other words, from the mathematical viewpoint the trivial is really trivial and yet, from the philosophical viewpoint, it is very interesting\*.

---

\* My kind thanks to Dr. James O'Hara (Hanover) for the translation and to Isolde Hein (Hanover) for additional help.

# Structure-Similarity as a Cornerstone of the Philosophy of Mathematics

IVOR GRATTAN-GUINNESS (Middlesex)

*How* does a mathematical statement mean in an empirical situation? Are the axioms of mechanics chosen for their epistemological role or for their empirical evidence? Does the algebra of logic have to reflect the laws of thought? How does a mathematico-empirical theory talk about the physical world? If scientific theories are guesses and may be wrong, then even *what* do they talk about, and what is the mathematics doing in there?

The best available answer to these questions is 'it depends'. This paper contains a variety of preliminary remarks in an attempt to get further. After some explanations in Part 1, a range of some case-studies is presented in Part 2 before proceeding to some general philosophy in Part 3. The notion of structure-similarity will be proposed as a fundamental component of this philosophy; set theory, logic and the normal "philosophy of mathematics" of today have a significant but restricted place.

## 1. Introduction

But I hope that I have helped to restart a discussion which for three centuries has been bogged down in preliminaries.

K.R. Popper [1972, vii]

### 1.1. Chains of Reference

My term 'structure-similarity' is, I believe, rather new, and its own similarities and differences from the more established categories need to be explained. The chief concern is with the *content* of a mathematical theory, especially the way in which its structure relates to that of other mathematical theories (which I call 'intramathematical similarity'), to that of a scientific theory to which it is on hire ('scientific similarity'), and to empirical interpretations of that scientific theory in reality ('ontological similarity'). When structure-similarity carries through from mathematics to science and on to reality, we have 'ontological similarity'; however, a skein of difficult questions arises there.



Scientific theories often reflect the character of our universe in containing transempirical components, so that it is not sufficient to specify simple correlations between theoretical components and directly empirical categories. When a particular similarity is *not* upheld, then either a reason is put forward to prevent its advocacy, or else some different structure may obtain.

Here are two very simple examples. The first concerns scientific similarity: if the integral is thought of an area or a sum, and  $\int f(x)dx$  says something about hydrodynamics, does it have to do so in areal or summatory terms? The second example uses empirical similarity: if  $a = b + c$ , and  $a$  is an intensity of sound, do  $b$  and  $c$  also have to be (added) intensities in order that the equation makes good acoustical theory? Both examples can be applied to intra-mathematical similarity also: if  $a$  and  $b$  are interpretable as lengths, say, or if the integral is so interpretable in a problem concerning conics.

The word 'applied' can itself be applied to all these examples, for it covers all three kinds of similarity: one can apply a mathematical theory a) elsewhere in mathematics, b) to scientific contexts, and c) as part of a mathematico-scientific theory, to reality. (For most purposes, the word 'applied' is perhaps *too* wide in its use, and one could argue for a return to the older (near) synonym 'mixed' to designate the second and third kinds.) In all kinds the possibility of *non*-similarity is to be borne forcibly in mind, and indeed the word 'similarity' is to be taken throughout the paper as carrying along its opposite also. The first, or intra-mathematical, kind of similarity has many familiar manifestations, and I shall not dwell too long on them in this paper, for my chief concern is with the other two kinds, which bear upon a major point of interaction between the philosophy of science and the philosophy of mathematics: the use of mathematics in scientific theories.

Various issues in the philosophy of science are involved. I leave most of them to one side here, referring to my paper [Grattan-Guinness 1986], to which this one is a sequel and a development. However, three features of that paper will be useful. Firstly, the remark that scientists 'reify' objects of concern when forming theories: they *suppose* that certain kinds of entity or process exist for the purpose of theory-building (and, for the concerns of this paper, may use mathematics in the process). Secondly, when testing theories they check if some of these reified entities and processes actually exist in reality: if so, then reification becomes reference. It may be that a full theory refers, in which case it becomes 'ontologically correct', a category I use to replace some of the misuses of truth. Finally, there is the notion of 'desimplification', in which a scientific theory is formulated in which various effects pertaining to the phenomena are *knowingly* set aside as negligible or at least too complicated to deal with in the current state of knowledge, but then

are reinstated later in desimplified forms of the theory, when the measure of scientific (and maybe also ontological) structure-similarity is thereby increased. I propose this notion for the philosophy of science as an improvement upon the category of ad hoc hypotheses, for there is no component of ad hocness in desimplification: one might even *hope* to show that the neglected effect is indeed small enough to be set aside in the measure of accuracy and fine detail within which the current scientific and experimental activity is set.

Our concern in this paper is with *how* a mathematical theory can mean, not *what* it may mean: my paper [Grattan-Guinness 1987b] is based upon the theme 'How it means' within a particular historical time-period. Here is one respect in which the differences of intent from much modern philosophy (of mathematics) are evident.

### *1.2. Plan and Purpose of the Paper*

Part 2 of the paper contains a selection of case-studies. After the intra-mathematical considerations of the next two sections, in which the main novelty is the distinction between icons and representations, scientific and empirical similarities dominate. Sections 2.2 and 2.3 contain some examples from the uses of the calculus in mechanics and mathematical physics, dealing in turn with the formation and the solution of differential equations. Some of them involve Fourier series, which also raise the question of linearisation of scientific theories; this matter is discussed in more general terms in section 2.5. Then in section 2.6 the focus turns to mathematical psychology, in the form of Boole's algebra of logic and its early criticism.

The examples used in Part 2 happen to be historical, and involve cases which I can explain without difficulty to the reader and with which I am sufficiently familiar to draw on with confidence for current purposes. But the points made are equally applicable to modern mathematical concerns, and I hope, therefore, that they will catch the interest of mathematicians as well as historians and philosophers.

In Part 3 more general and philosophical considerations are presented. Section 3.1 contains remarks on the limitations of axiomatised mathematical theories. In the same spirit, doubts are expressed in section 3.2 about philosophers' 'philosophy of mathematics', which is centered on logics and set theories but rarely get further; and mathematicians' 'philosophy of mathematics', where the breadth of the subject to noted (to some extent) but

logical theory and related topics are disregarded. An alternative philosophy is outlined in sections 3.3 – 3.5, in which both the range of the subject and logic (and related topics) are taken seriously.

## 2. Some Case-Studies

### 2.1. Intra-mathematical Similarity Between Algebra and Geometry

It is a commonplace that different branches of mathematics relate by similarity to each other. The example of  $a + b$  as numbers and as lengths, just mentioned, is a canonical example of the numerous cases which apply to the complicated relationships between geometries and algebras: my use of plurals here is deliberate. It can be extended to the possibility of non-similarity, for at various times objections have been made to the legitimacy of negative numbers [Pycior 1987], so that their interpretation as suitably directed line segments was not adopted. Other such cases include the interpretation of powers of variables relative to spatial dimensions: it is striking to note that when [Viète 1591] advocated the new ‘analytic art’ of algebra he used expressions such as ‘squared-cube’ to refer to the fifth power, with similar locations for all powers above the third, and thereby to draw on structure-similarity between algebra and geometry, and (by an implication which involves a huge burden of questions concerning our theme), onto space. By contrast, a few decades later [Descartes 1637] showed no qualms in his *Géométrie* when advocating higher powers in his algebraisation of geometry, and in writing  $z^4$  just like  $z^3$ , and thereby discarding this similarity; in the same way he regarded negative roots of equations as ‘false’ and so dispensed with that possible link also.

In some respects, therefore, Descartes was a non-similarist. Yet at the beginning of Book 3 he announced that ‘all curved lines, which can be described by any regular motion, should be received in Geometry’, a typical example of an isomorphism between one branch of mathematics and another. Notice, however, that the similarity often does not go too far. For example, there are no obvious structural similarities between basic types of algebraic expressions and fundamental types of geometrical curve; hence classification has been a difficult task for algebraic geometry (or should it sometimes be thought of as geometrical algebra?) from Newton onwards. Again, dissimilarity is evident in problems involving the roots of equations, where the algebra is unproblematic but a root does not lead to a geometrically intelligible situation (an area becomes negative, say). The occurrence of complex zeroes in a polynomial can be still more problematic as geometry, since their presence is not reflected in the geometrical representation of the

corresponding function in the real-variable plane, and a pair of complex-variable planes will represent the argument and the value of the function but necessarily do not reflect the function itself. Other cases includes a situation where the algebraic solution of a geometrical problem supplies a circle (say) as the required locus whereas only an arc of it actually pertains to the problem.

In an undeservedly forgotten examination of 'the origin and the limits of the correspondence' between these two branches of mathematics. [Cournot 1847] explores in a systematic way these and other cases and sources of such structural non-similarity. The mathematics is quite elementary; the philosophy is far from trivial.

## *2.2. Representations and Icons*

Sometimes these intra-mathematical structure-similarities are put forward more formally as representations, such as in the geometrical characterisation of complex numbers in the plane, or Cauchy's definition by such means of infinitesimals in terms of sequences of real values passing to zero. There is no basic distinction of type in these cases, but considerable variation in the manner and detail in which the structure-similarity or representation is worked out: Descartes' case was to cause him and many of his successors great difficulty concerning not only over the detail but also the generality of the translation that was effected. In section 3.1 we shall note an approach to mechanics in which it was explicitly avoided.

A particular kind of intra-mathematical structure-similarity worth emphasising is one in which the mathematical notation itself plays a role, and is even one of the objects of study. Within algebra matrices and determinants are an important example, in that the array can be subjected to analysis (by graph-theoretic means for large sparse matrices, for example). Following C.S. Peirce, I call them 'icons', and draw attention to one of his examples, in algebraic logic, where systems of connectives were set up in squares and other patterns in ways which reflect their significations [Zellweger 1982].

Another type of example is shown by algebraic logic: the principle of duality, which was exploited by Peirce's contemporary E. Schröder. He stated theorems in pairs, deploying a formal set of (structure-conserving) rules of transformation of connectives and quantifiers in order to get from one theorem to the other. He consciously followed the practise of J.V. Poncelet and J.D. Gergonne, who had laid out theorem-pairs in projective geometry following a set of rules about going from points, lines and planes. A current interest in logic is a type of generalisation of duality into analyses of proof-structures in order to find structure-isomorphisms between proofs and maybe to classify

mathematical proofs into some basic structural categories. In cases like these, the mathematical text itself is the/a object of study, not (only) the referents of the mathematical theory. The situation is not unlike the use of a laboratory notebook in empirical science, where the matters of concern can switch back and forth between the laboratory experiments and the contents of the notebook itself.

### *2.3. Modelling Continua: the Differential and Integral Calculus*

The calculus has long been a staple method in applied mathematics, especially in the differential and integral form introduced by Leibniz. Here the principal device was represent the (supposed) continuity of space and time by the theory of differentials, infinitesimally small forward increments  $dx$  on the variable, with its own second-order infinitesimals  $ddx$ , and so on [Grattan-Guinness 1990a]. (Newton presented his second law of motion under the same regime, in that he stated it as the successive action of infinitesimally small impulses). The key to the technique lay not only in the smallness of the increment (a controversial issue of reification, of course) but especially in the preservation of dimension under the process of differentiation: if  $x$  is a line, then so is  $dx$  (a very short line, but infinitely longer than  $ddx$ ). The various orders of infinitesimal were used to reflect individually the increments on the variables of the chosen problem, and literally 'differential equations' were formed: that is, equations in which differentials were related according to some physical law. The measure of structure-similarity is quite substantial: for example, a rate of change  $dx/dt$  was the ratio  $dx+dt$  of two infinitesimal increments, a property lost in an approach such as Cauchy's or Newton's based on limits, where the derivative  $x'(t)$  does not reflect its referent in the same way.

The integral branch of the calculus also can exhibit issues pertaining to our theme when interpreted as an area or as an infinitesimal sum. In energy mechanics, for example, the work function  $\int Pds$  designates the sum of products of force  $P$  by infinitesimal distance  $ds$  of traction, and founders of this approach in the 1820s, such as Poncelet and G.G. Coriolis, explicitly made the point. Before them, P.S. Laplace had inaugurated in the 1800s a programme to extend the principles of mechanics (including the use of the calculus) to the then rather backward discipline of physics by modelling "all" physical phenomena on cumulative intermolecular forces of attraction and repulsion [Grattan-Guinness 1987a]. The key to his mathematical method was to reflect the cumulative actions by integrals; but his follower S.D. Poisson modified this

approach in the 1820s by representing the actions instead by sums. Cauchy did the same in his elasticity theory of the same time, and for a clearer reason than Poisson offered, more clearly involving structure-similarity: the magnitude of the action on a molecule was very sensitive to locations of the immediately neighbouring molecules, and the integral would not recognise this fact with sufficient refinement [Grattan-Guinness 1990b, 1003–1025]. [Cournot 1847] is again worth consulting (at chs. 13–15), for a range of examples concerning not only the basic aspects of the calculus but also rectification and the definability of functions as definite integrals.

#### 2.4. *Linearity or Non-linearity: the Case of Fourier Series*

The elegant body of mathematical theory pertaining to linear systems (Fourier analysis, orthogonal functions, and so on), and its successful application to many fundamentally linear problems in the physical sciences, tends to dominate even moderately advanced University courses in mathematics and theoretical physics. [... But] nonlinear systems are surely the rule, not the exception, outside the physical sciences.

May [1976], 467

Let us take now a major example of ontological similarity. One of the great phases which led to the great importance of linear modelling of this non-linear world was linked with the rise of classical mathematical physics on the early years of the 19th century. The note just taken of the elasticity theory of that time was part of this adventure, and Cauchy is prominent example of a linearist. Another major figure was Fourier, who introduced the mathematical theory of heat diffusion and radiation from the 1800s onwards. His deployment of linearity embraced not only the assumptions used in forming the diffusion equation but also in solving it by Fourier series.

These series form an excellent example of ontological non-similarity and incorrectness: not only the (non-)similarity of these linear solutions with the phenomena to which they refer but also the question of their manner of representing (in Fourier's case) heat diffusion. Mathematically speaking, they comprise a series of terms exhibiting integral multiples of a certain periodicity specified by the first term, sometimes prefaced by a constant term; in addition, sine/cosine series exhibit evenness/oddness. They describe diffusion at the initiation of the time  $t$ : for later values of  $t$  exponential decay terms are multiplied into the time terms.

Now the ontological structure-similarity of these time terms is clear (and indeed rather important for the legitimacy of the analysis, especially for its non-linear critics!); but the trigonometric terms raise questions. Does heat

have to be interpreted as waves, in families of corresponding periodicities? The nature of heat was an important question at that time, with debates as to whether it was a substance (caloric), a product of molecular action, an effect of vibratory motion (that is, the waval theory just mentioned), or something else; but Fourier himself did not like the question, preferring to treat heat and cold as opposites without reifying their intimate structure. His most explicit approach to a structure-similar reading occurred in a treatment of the solution of Laplace's equation (as the case of the diffusion equation for steady state in the lamina in the Oxy plane):

$$(1) \quad \sum_{r=0}^{\infty} a_r \cos ry e^{-ry} \quad ;$$

here he spoke of each component term 'constitut[ing] a proper and elementary mode', such that 'there are as many different solid laminae that enter in to the terms of the general surface, that each of these tables [laminae?] is heated separately in the same manner as if there were only a single term in the equation, that all these laminae are superposed' [Fourier 1807, 144]. Even here he went so far only to affirm superposition; the periodic character of the functions involved was not affirmed ontologically as heat behaviour.

Fourier series are well known also in acoustics, and indeed they had already been proposed in this context, especially by Daniel Bernoulli. Here structure-similarity between periodicities and pitch level were proposed, together with further similarities with pendular motion [Cannon and Dostrovsky 1981]: 'for the sounds of horn, trumpets and traverse flutes follow this same progression 1,2,3,4,..., but the progression is different for other bodies' [Bernoulli 1755, art. 3]. He did not have the formulae for the coefficients of the series; by contrast, Fourier did know them, and also the manner of representing the function by the series outside its period of definition, but when discussing his predecessors on the matter he did not follow Bernoulli's advocacy of structure-similarity when vindicating his case against the criticisms offered at the time by Euler and d'Alembert.

Bernoulli's stance was adopted much later by G.S. Ohm, a German scientist not so oriented towards Fourieran positivism (although much influenced by Fourier's methods in his earlier work on electromagnetism). In his paper [Ohm 1843] he rejected the usual view of sound as composed of a small number of simple components and proposed instead that (infinite) Fourier series could be so structurally interpreted: 'If we now represent by  $F^1$  any sound impulse striking the ear at time  $t$ , then Fourier's theorem says that this impulse is analysable into the [trigonometric] components', where the constant term 'corresponds to no oscillation but represents merely a displacement of the

oscillating parts of the ear. The other components, however, all correspond to oscillations which take place around the displaced position'.

In this way Ohm imposed the structure-similarity of Fourier series onto acoustic theory in a way which Bernoulli had envisaged a century earlier but without the details of the mathematical theory which Fourier was to provide between them. From then on this interpretation of series became well-known, and even of mechanical representation in Kelvin's harmonic analyser. Kelvin is a wonderful subject for our theme, for he was an advocate of the method of analogy from one theory to another; for example, his early work in electro-magnetism was based on a 'flow analogy' from Fourier's theory of heat [Wise 1981]. His machine evaluated integrals of the form

$$(2) \quad \int f(x) \frac{\sin}{\cos} nx \, dx \quad \text{for } 0 \leq n \leq 2$$

via the motion of literally rotating cranks to produce the trigonometric components. In the particular cases of tidal theory, the terms formed which corresponded to the various 'tidal constituents' held to be created by the various solar and lunar effects; the formula was also applicable to the important problem of compass deviation on iron ships [Kelvin 1882]. Later workers developed de-simplified versions of this machine which could calculate the Fourier coefficients for arbitrarily large values of  $n$ , for a variety of purposes [Henrici 1892].

Here we see scientific structure-similarity in a stark form: the tidal level was calculated by Kelvin as literally the sum of various components. Yet in this application at the same time *non*-similarity from another mathematical point of view is also evident, rather like Fourier's advocacy of superposition but without commitment to a waval theory of heat; for the trigonometric terms as such were not interpreted as a planetary effect.

This tradition goes back to Euler's celestial mechanics, when he took Newton's second law to express both planetary/solar and inter-planetary action; the latter group of effects involved powers of the appropriate distance functions, which were stated via the triangle formula. Now this expression took the form  $(a + b \cos \alpha)^{-3/2}$ , where  $a$  was the angle between the radius vectors of the two planets involved; and to render this expression in more amenable form he used De Moivre's theorem to convert it into a series of the form  $\sum_r (a_r \cos r\alpha)$  [see, for example, Euler 1749]. But this is trigonometric series once again, in a different context. However, this time they arose out of a pure-mathematical artifice, *bereft* of the structural interpretation that was then being advocated by Bernoulli for acoustical theory (and being *rejected* by the same Euler in favour of functional solution of the wave equation).



## 2.5. *Linearity or Non-linearity: the General Issue*

I remember being quite frightened with the idea of presenting for the first time to such a distinguished audience the concept of linear programming.

After my talk the chairman called for discussion. [... Hotelling] said devastatingly: '*but we all know the world is non-linear*'. [...]

Suddenly another hand in the audience was raised. It was von Neumann. [...] 'The speaker called his talk "Linear programming". Then he carefully stated his axioms. If you have an application that satisfies the axioms, use it. If it does not, then *don't* ', and he sat down.

Dantzig [1982, 46]

If scientific and ontological structure-(non)-similarity are not incorporated into the philosophical scenario of developments such as the ones described in the previous section, then the full richness of the intellectual issues that were at stake cannot be appreciated. And the concerns are not only historical: indeed, the rise of computers and the new levels of efficacy achievable in non-linear mathematics has raised the issue of linear versus non-linear reification and reference to a new level of significance. Acoustics and related branches of science themselves form a fine example of this trend, for especially since the 1920s attention has focussed on non-linear oscillations of all kinds, especially the so-called 'relaxed' variety, and upon associated phenomena known as 'irregular noise' some decades ago and now carrying the trendy name 'chaos' (see the survey in [West 1985, ch. 3], a profound study of 'the importance of being nonlinear').

Competition between utility of linear or non-linear versions of theories in same range of concern can ensue: an interesting example is linear and non-linear programming, in which the latter was instituted in the early 1950s soon after the establishment of the former, but took rather different origins [Grattan-Guinness 1992a]. Hotelling's reservation about linear programming, quoted at the head of this section, was soon to be dealt with, at least in part.

From the point of view of the philosophy of mathematics, the extension of structure-similarity is a central feature of the questions raised in this and the previous sections. For the philosophy of science, desimplification of theories, and the realm of legitimate reification, are correspondingly central concerns. For science itself, the manner of extending theory from mechanics and physics to other branches is a major component issue.

## 2.6. *Mathematical Psychology and the Algebra of Thought*

My last context seems to be quite different; yet we shall meet some striking

metasimilarities. It concerns Boole's formulation of an algebra to represent normatively 'The laws of thought' (to quote the title of his second book on the matter, of 1854). (This was done before Peirce's concerns with semiotics mentioned in section 2.2.) The basic law  $x^2 = x$  of 'duality' obeyed by 'elective symbols'  $x$  (which selected the individuals satisfying a property such as 'European') differentiated the algebra of his logic from the common algebras, which were structurally inspired by arithmetic; but he maintained similarity by deploying the other three connectives, and made use of subtraction, division and addition as well as the multiplication involved in the law of duality.

The chief object of concern here is Boole's definition of '+'; for convenience I shall use his interpretation in terms of classes.  $(x+y)$  was defined so that 'the symbol +' was 'the equivalent of the conjunctions "and", "or"': however, 'the expression  $x + y$  seems indeed uninterpretable, unless it be assumed that the things represented by  $x$  and the things represented by  $y$  are entirely separate; that they embrace no individual in common' [1854, 55, 66]. The specification of the unions of non-disjoint classes  $x$  and  $y$ , corresponding to the exclusive and inclusive senses of 'or', required the intermediate definitions of disjoint classes:

$$(3) \quad x(1 - y) + y(1 - x) \text{ or } x + y(1 - x) ,$$

as required (p. 57).

Thus we see that the symbol ' $x + y$ ' actually reflected a process structurally similar to addition, although it was defined under the hypothesis of the disjointness of the components. Boole tried to argue for the necessity of this restriction; but his arguments were hardly convincing, and they were rejected by his first major commentator, W.S. Jevons, in his little book *Pure logic* [Jevons 1864]. His subtitle, 'the logic of quality apart from quantity', expressed his desire to reduce the structural links with mathematics as espoused by Boole (although this and other work by Boole himself and others showed that the distinction between quality and quantity did not capture the essentials of the mathematics of that time). Jevons worked with 'terms', that is, 'any combination of names and words describing the qualities and circumstances of a thing' [1864, 26, 6]; a proposition '*is a statement of the sameness or difference of meaning between two terms*', to be written  $A = B$  with '=' read as 'is' (pp. 8-9). He followed Boole in 'combining' terms  $A$  and  $B$  in a Boolean manner to produce  $AB$ , and accepted the law of duality (as a law satisfied by terms, giving it the name 'the law of simplicity'); but he rejected entirely the restriction laid upon the definition of addition, arguing that the natural use of language permitted the definition of full union of

intersecting classes (or overlapping terms). '*B or C* is a plural term [...] for its meaning is either that of *B* or that of *C*, but it is not known which' (p. 25: later examples show that inclusive disjunction was intended). His 'law of unity' for a term *A* allowed that

$$(4) \quad A + A = A ;$$

that is, two (and therefore any number) of self-disjunctions of *A* may be reduced to a single *A* without change of meaning. Hence logical alternation was different from mathematical addition: intra-mathematical structure-similarity was being rejected. Jevons distanced himself further from Boolism by dispensing entirely with division and subtraction.

Boole and Jevons corresponded on these matters in 1863–1864, some months before Boole's unexpected death [Grattan-Guinness 1991a]. They took as a "test case" the example of  $(x + x)$ . For Jevons it satisfied (4); for Boole it was not interpretable at all; although in his system it followed that the equation  $x + x = 0$  was reducible to  $x = 0$  via one of his general expansion theorems for a general logical function. Thus the differences in ontological structure-similarity ran quite deeply, and Jevons found Boole to be ontologically incorrect.

These changes required Jevons to develop different methods of obtaining consequences (or, as he called them, 'inferences'). 'Direct inference' worked in effect on the transitivity of terms equated in the premises; for example, from  $A = B$  and  $B = C$ ,  $A = C$  could be inferred [Jevons 1864, 10–13]. In the more powerful method of 'indirect inference' he formed all possible logical combinations of the simple terms involved in the premises, together with the contrary terms, combined these compound terms with both members of each premise, and retained only those terms which either were consistent with both members of at least one premise or contradicted both members of all of them. The consequences were drawn by taking any simple or compound term (*C*, say) and equating it to the sum (in his sense of '+') of all the retained terms of which it was part. In other words, he found the plural term to which *C* was equal ('=') under the premises. Basic laws such as duality and simplicity, and various rules of elimination, simplified the resulting propositions (pp. 42–53).

This procedure of selecting and inspecting was rather tedious: in order to render it more efficient [Jevons 1870] introduced his 'logical machine'. The similarity of function with the harmonic analysers of Kelvin and his successors in section 2.5 is worth noting: in both cases a structure was carried through from a theory and its referents to a mechanical imitation.

### 3. *Towards a General Philosophy of Mathematics*

If I speak very decidedly about the consequences of the neglect of pure logic by mathematicians, as I have done elsewhere about the neglect of mathematical thought by logicians, I shall not be supposed to have any disrespectful intention [...]

Augustus De Morgan [1869, 180], alluding to George Peacock

#### 3.1. *Prologue: the Limitations of Axiomatisation*

I must elaborate on the matter of axiomatisation. Under a deeply unfortunate educational tendency, itself partly inspired by the so-called ‘philosophy of mathematics’, many theories are formulated in a neo-axiomatic way; and the impression accrues that the subject is deeply unmotivated. In many cases this impression arises because the axioms are used precisely because of their epistemological status, *as* starting-point for deductions: they have little or no intuitive character (which is why they are so often hard to find or identify as axioms in the first place!). In particular, structure-similarity is rarely evident: even the intuitive feel in Boole’s algebra is largely lost in the axiomatisation of the propositional calculus (which is not due to Boole or Jevons, incidentally). Further, the axiomatised states have nothing to say about the desimplification of theories – ironically, not even about the enrichment of axiomatised versions! – and thus do not focus upon scientific knowledge as a process of growth.

Take classical mechanics, which was and is an especially rich branch of mathematics from our point of view, with its variety of formulations [Grattan-Guinness 1990c]. The most axiomatised version is the variational tradition, of which two main forms developed: Lagrange’s, based on the principles of virtual work and of least action; and the extended version based on Hamiltonians. In both cases the aim was to develop a very general theory from few assumptions: the aim of those times corresponding (but *not* in close detail) to our conception of axiomatisation. But, as in modern cases, the price of (alleged) generality is intuition: one cannot claim the Lagrange equations as an *evident* way of founding dynamics. The reason can be stated in terms of structure-non-similarity: the various terms do not imitate in any way the phenomena associated with their referents, and neither do their sums or differences.

Allied with variational mechanics is potential theory, which has a curiously ambiguous place in our context. On the one hand some potentials, such as the velocity potential, for example, do not have a clear referent (so that structure-similarity is ruled out); and the point is important from an epistemo-

logical angle, since the potential of some category  $X$  could replace  $X$  in the list of ontological commitments (for example, if the work expression always admits a potential, does 'force' become only *façon de parler*?). However, equipotential surfaces can, and indeed are intended to, take a direct reification and even reference – as stream lines in hydrodynamics, or as the defining surfaces for lines of optimal electromagnetic flow.

### 3.2. *Between Two Traditions*

It is clear that by 'mathematics' I intend to refer to "all" of the subject, both ancient and modern, and also the modes of reasoning which attend it. Thus I try to bridge the following lamentable gap.

On the one hand, there is the "philosophy of mathematics" which starts out from logics, set theories and the axiomatisation of theories, but rarely gets much further. It has flourished since the 1920s, mostly in the hands of philosophers. Without doubt important and fruitful insights have come out of this tradition, often with consequences beyond the fragments of mathematics studied; but they belong only to a corner of the wide range of questions which the philosophy of actual mathematics excites. One might as well think that music is the same as piano sonatas.

On the other hand, there is the opposite absurdity practised by mathematicians, which respects (much of the) range of their subject but tends to adopt the metaphilosophy of ignoring the logico-philosophy of the subject. It takes logics and proof theory for granted even if axiomatisation is (over)-emphasised; for the logical issues at hand are often poorly understood [Corcoran 1973].

This separation has long been in place, unfortunately, in one form or another. In a paper [Grattan-Guinness 1988] I surveyed the contacts between logics and mathematics between the French Revolution and the First World War, and I gave it the title *Living together and living apart* to underline the modest degree to which contacts functioned.

Both traditions are quite often formalist in character, either in the technical sense of the word associated with Hilbert or in a more general way as unconcerned with reference [compare Goodman 1979]. Further, while structure-(non)-similarity is a major component, this philosophy has little purpose in common with the structuralist philosophies, which are often involved only with set-theoretic formulations and/or abstract axiomatisations of theories, with associated model theory [Vercelloni 1988]. It is time to bring these two traditions together, in philosophy which takes 'how it means' as the prime question rather than the 'what it means?' of normal modern philosophy (of

mathematics). In the next three sections I shall show how *both* philosophies can and must be accommodated.

### 3.3. *The Philosophy of Forms*

There is nothing at all new in my emphasis on the use of one theory ( $T_1$ , say) in another one ( $T_2$ ). But new to me, anyway, is the importance of structure-*non*-similarity between theories, and the consequent idea of *levels of content* which  $T_2$  may (or may not) exhibit in  $T_1$ . From this start I propose the following formulation, in which intra-mathematical and scientific structure-similarity (and non-similarity) are borne in mind.

Mathematics contains *forms*, which may be expressions, equations, inequalities, diagrams, theorems with proofs, even whole theories. Some forms are atomic, and can be concatenated together to produce *superforms*, or *compound forms*. The level of atomicity can be varied, depending on the need and context: thus in hydrodynamics the integral may be used as an atom, but in the foundations of analysis it would be dissected into its components. (An example for this case is given in section 3.5 below.) The forms themselves characterise mathematics, and distinguish it from, say, chemistry.

The assembly of forms is very large. It includes not only those from abstract algebra such as group and field but also integral (as just mentioned), exact differential, least-squares, addition, neighbourhood, limit, and so on and on. To each form there are sub- or special forms: double integrals and Abelian groups, for example. An important special case of intra-mathematical structure-similarity was stressed in section 2.2 under 'icons': mathematical notations, which themselves can play the role of  $T_2$ .

When  $T_1$  is applied to  $T_2$ , the repertoire of forms of  $T_1$  work in  $T_2$ , but with differing levels of content (some clear examples were given in Part 2). Thus one can understand *how* the 'unreasonable effectiveness of mathematics in the natural sciences' occurs: there is no need to share the perplexity of [Wigner 1960] on this point if one looks carefully to see what is happening. Further, the notion of desimplifying scientific theories can be used here; for two of its sources are the increase in structure-similarity and in levels of content. The *genuine* source of perplexity that the mathematician-philosopher should consider is the *variety* of structures and of levels of content that can obtain within one mathematico-scientific context.

### 3.4. *The Philosophy of Reasonings and Structures*

So far no notice has been taken of logics and allied theories. I bracket them

together under the word 'reasonings'. This word is chosen without enthusiasm, but every candidate synonym or neonym is defective: 'logic(s)', 'deduction' and 'proof' are definitely too narrow, and 'argument' and 'connection' not sufficiently specific. In this category I place logics, both formal/axiomatic and natural deduction, bivalent and also non-classical forms, including predicate calculi with quantification and set theories; the associated metatheories are also on hand. In addition, place is granted to rules for valid and invalid inference, and the logic of necessary and sufficient conditions for the truth of theorems; and for proof methods such as mathematical induction, by reduction to the absurd, by modus ponens, and so on, again with special kinds (first- and higher-order induction, for example). There is quite a bit of repetition in this catalogue, for many of these methods could be expressed within some logical frameworks; but the agglomeration is not offensive to this sketch.

In addition, we need definitional theory, a neglected area [Gabriel 1972] where mathematics often comes adrift [Dugas 1940; Rostand 1960]. Topics include rules of well-formation of nominal definitions in formal but also non-formal mathematical theories; formation to meet given criteria (such as the correlation coefficient in statistical regression, or Yule's structurally similar Q-parameter for association); definitional systems and their relationship to axioms; and creative and contextual definitions. Some of these types of definition involve model-theoretic notions such as (non)-categoricity [Corcoran 1980], so that they sit here also. In addition, philosophical ideas about existence and uniqueness of defined terms will need consideration. Finally, some philosophy of semiotics will be needed to appraise the use of iconic forms.

The main distinction between forms and reasonings is that the former pertain to mathematics itself while the latter are suitable in other areas of thought. But they have an important common factor: like forms, reasonings have structures, which are *not* objects in the way that forms and reasonings are (at least, as they are in the liberal ontology which I have admitted into this sketch). Take the example mentioned several times in Part 2 of the integral, treated as a superform under the structural concatenation of the forms limit, function, sum, difference and product, together with the reasonings of nominal definition (':=') and existence:

$$(5) \quad \int f(x)dx := \lim \sum_r [f(x_r) \Delta x_r] \text{ as } \Delta x_r \rightarrow 0 \text{ if the limit exists.}$$

The structure *glues this concatenation together*, but it is not itself an object of the concatenation. Further, this structure is distinguished from other structures by the glueing, which tells the integral as this sum apart from the integral as an infinitesimal sum or as the inverse of a derivative. An analogy may be

drawn between a book, any of its chapters, and its price or weight: the book and the chapters are objects in a sense denied to the price and the weight.

### 3.5. *Mathematics as Forms, Reasonings and Structures*

Several features of mathematical development and progress can be illuminated by the philosophy proposed here. The important notion of analogy [Knobloch 1989] falls into place:  $T_1$  has been applied with success to  $T_2$ , and task of analogy is to assess the similarity of structure between  $T_1$  in  $T_2$  and  $T_1$  in some new context  $T_3$  (as Kelvin did with the success noted in section 2.4). Again, the process of ‘having a new idea’ and thereby advancing mathematical knowledge can be given a more precise characterisation in those cases (which will be the vast majority) in which the idea is *not* completely novel in itself. The famous phrase of the novelist E.M. Forster, ‘only connect’, is apposite: *new* combinations of forms, structures and reasonings are made, and a theory developed from there. Similarly, missed opportunities are situations where the connections are not made [Dyson 1972].

This philosophy also has the advantage of conveying the great multitude of complications, of all kinds, that attend mathematics. As was mentioned in section 3.2, the number of forms is very great, and a rich basket of reasonings is involved also; thus the assembly of structures is commensurately large. In addition, the chains of reference (section 1.1) are varied in their kinds: from mathematics to scientific theory and maybe on to reality. Finally, there is a range of possibilities in a theory for structure-similarity to be upheld between some components but denied to others (for example from section 2.4, Fourier on superposition but not on the waval interpretation of heat).

When this philosophy is used historically (as just now), the usual caveats about anachronism would have to be watched with especial care; in particular, the ignorance of logic among mathematicians will require the historian to deploy reasonings with especial delicacy. But, as the examples of Part 2 show, this philosophy has much to say about the past. Here is another important source of difference of this philosophy from traditions which take no notice of the evolution or development of a theory: for example, here is a further manner of expressing the reservations of section 3.1 on exaggerating the place of axiomatisation. I am much more in sympathy with the view of [Polya 1954 and 1962-1965] and of followers such as [Lakatos 1976], and also the unfairly neglected [Rostand 1962], on the role of modification of proofs to generate mathematics – which, among other things, is an historical process. A few other modern philosophies of mathematics take history seriously in some way [see *passim* in Tymoczko 1985].



When one takes on also the use of probability and statistics within this context, a range of *quite basic* additional issues concerning purpose of theory are raised: probability as a compensation for ignorance or as a genuine category for reference or reification, and the various interpretations of probability and their referentiability. It is a great pity that questions of these types do not occupy an important place in the practise of the philosophy of mathematics, for they occupy comparable positions in mathematics itself.

Yet even now not every question involved in the philosophy of mathematics has been raised. For example, the creative side of the subject is not basically touched, although, for example, the questions concerning mathematical heuristics may be tackled from the point of view of maximising levels of content among the various presentations available [compare Polya 1954]. As a special case this remark restates once again my criticisms of axiomatisation made in section 3.1.

To conclude, in this philosophy mathematics is seen as a group of problems, topics and branches in which forms and reasonings are chosen and deployed in a variety of structures exhibiting differing levels of content. This, briefly, is *how* this philosophy of mathematics means, and also how this philosophy of mathematics *means*.

A developed version of would be highly taxonomic in character. What are the atomic forms, the reasonings, and structures? How do they relate to each other; via (meta)structural isomorphism for example? Does it matter that versions of set theory occur both as forms and in reasonings? What, if anywhere, is the place of a priori knowledge? I am not at all sure of the answers to these questions; but I am sure that they are *fruitful* questions, and examining them could close the lamentable gap that exists between the practise of mathematics and the reflections upon it that mathematicians and philosophers make. It is a great pity that the philosophy of mathematics is always bogged down in preliminaries.

### *Bibliography*

- Bernoulli, D. 1755. "Réflexions ... sur les nouvelles vibrations des cordes ...", *Mémoires de l'Académie des Sciences de Berlin*, 9 (1753: publ. 1755), 147-172.
- Boole, G. 1854. *The laws of thought* ..., London (Walton and Maberley). [Repr. 1958, New York (Dover).]
- Cannon, J. T. / Dostrovsky, S. 1981. *The evolution of dynamics: vibration theory from 1687 to 1742*, Heidelberg (Springer).
- Corcoran, J. 1973. "Gaps between logical theory and mathematical practise". In: M. Bunge (ed.), *The methodological unity of science*, Dordrecht (Reidel), 23-50.

- Corcoran, J. 1980. "Categoricity", *History and philosophy of logic*, 1, 187-207.
- Cournot, A. A. 1847. *De l'origine et des limites de la correspondance entre l'algèbre et la géométrie*, Paris and Algiers (Hachette).
- Dantzig, G. B. 1982. "Reminiscences about the origins of linear programming", *Operations research letters*, 1, 43-48. [Slightly revised version in A. Bachem, M. Grottschel and B. Corte (eds.), *Mathematical programming. The state of the art*, 1983, Berlin (Springer), 78-86.]
- De Morgan, A. 1869. "On infinity and the sign of equality", *Transactions of the Cambridge Philosophical Society*, 11 (1864-69), 145-189.
- Descartes, R. 1637. "La géométrie", In *Discours de la méthode...*, Leiden, 297-413. [Various reprints and translations.]
- Dugas, R. 1940. *Essai sur l'incompréhension mathématique*, Paris (Vuibert).
- Dyson, F. 1972. "Missed opportunities", *Bulletin of the American Mathematical Society*, 78, 635-652.
- Euler, L. 1749. *Recherches sur la question des inégalités du mouvement de Saturne et de Jupiter*, Paris (Académie des Sciences ). Also in *Opera omnia*, ser. 2, vol 25, 45-157.
- Fourier, J. B.J. 1807 "Mémoire sur la propagation de la chaleur", ms. in *ENPC*, ms. 1851. Published in [Grattan-Guinness and Ravetz 1972] *passim*.
- Goodman, N. 1979. "Mathematics as an objective science", *American mathematical monthly*, 86, 540-551.
- Gabriel, G. 1972. *Definitionen und Interessen. Über die praktischen Grundlagen der Definitionenlehre*, Stuttgart (Frommann).
- Grattan-Guinness, I. 1986. "What do theories talk about? A critique of Popperian fallibilism, with especial reference to ontology", *Fundamenta scientiae*, 7, 177-221.
- Grattan-Guinness, I. 1987a. "From Laplacian physics to mathematical physics, 1805-1826". In: C. Burrichter, R. Inhetveen and R. Köter (eds.), *Zum Wandel des Naturverständnisses*, Paderborn (Schöningh), 11-34.
- Grattan-Guinness, I. 1987b. "How it means: mathematical theories in physical theories. With examples from French mathematical physics of the early 19th century", *Rendiconti dell'Accademia del XL*, (5)9, pt.2(1985: publ. 1987), 89-119.
- Grattan-Guinness, I. 1988. "Living together and living apart: on the interactions between mathematics and logics from the French Revolution to the First World War", *South African journal of philosophy*, 7, no. 2, 73-82.
- Grattan-Guinness, I. 1990a. "Small talk in Parisian circles, 1800-1830: mathematical models of continuous matter". In: G. König (ed.), *Konzepte des mathematischen Unendlichen*, Göttingen (Vandenhoeck und Ruprecht), 47-63.
- Grattan-Guinness, I. 1990b. *Convolutions in French mathematics, 1800-1840. From the calculus and mechanics to mathematical analysis and mathematical physics*, 3 vols., Basel (Birkhäuser) and Berlin (Deutscher Verlag der Wissenschaften).
- Grattan-Guinness, I. 1990c. "The varieties of mechanics by 1800", *Historia mathematica*, 17, 313-338.
- Grattan-Guinness, I. 1991a. "The correspondence between George Boole and Stanley Jevons, 1863-1864", *History and philosophy of logic*, 12, 15-35.

- Grattan-Guinness, I. 1992a. "'A new type of question': on the prehistory of linear and non-linear programming, 1770-1940", in E. Knobloch and D. Rowe (eds.), *History of modern mathematics*, vol. 3, Boston (Academic Press), to appear.
- Grattan-Guinness, I. with the collaboration of Ravetz, J.R. 1972. *Joseph Fourier 1768-1830. A survey of his life and work, based on a critical edition of his monograph on the propagation of heat, presented to the Institut de France in 1807*, Cambridge, Mass. (MIT Press).
- Henrici, O. 1892. "Über Instrumente zur harmonischen Analyse". In: W. von Dyck (ed.), *Deutsche Mathematiker-Vereinigung. Katalog mathematischer und mathematisch-physikalischer Modelle, Apparate und Instrumente*, Munich (Wolf), pt.1, 125-136.
- Jevons, W. S. 1864. *Pure logic, or the logic of quality apart from quantity*, London (Stanford). Also in his [1890], 1-77.
- Jevons, W. S. 1870. "On the mechanical performance of logical inference", *Philosophical transactions of the Royal Society of London*, 160, 497-518. Also in his [1890], 137-172.
- Jevons, W. S. 1890. *Pure logic and other writings* (ed. R. Adamson and H.A. Jevons), London (Macmillan).
- Kelvin, Lord 1882. "The tide gauge, tidal harmonic analyser, and tide predictor", *Minutes of the proceedings of the Institute of Civil Engineers*. Also in *Mathematical and physical papers*, vol. 6 (ed. J. Larmor), 1911, Cambridge UP, 272-305.
- Knobloch, E. 1989. "Analogie und mathematisches Denken", *Berichte zur Wissenschaftsgeschichte*, 12, 35-47.
- Lakatos, I. 1976. *Proofs and refutations ...*, Cambridge UP.
- May, R. M. 1976. "Simple mathematical models with very complicated dynamics", *Nature*, 261, 459-467.
- Ohm, G. S. 1843. "Über die Definition des Tons...", *Annalen der Chemie und Physik*, 59, 513-566. Also in *Gesammelte Abhandlungen* (ed. E. Lommel), 1911, Leipzig (Teubner), 587-633.
- Polya, G. 1954. *Mathematics and plausible reasoning*, 2 vols., Oxford (Clarendon Press).
- Polya, G. 1962-1965. *Mathematical discovery*, New York (Wiley).
- Popper, K. R. 1972. *Objective knowledge*, Oxford (Clarendon Press).
- Pycior, H. 1987. "British abstract algebra: development and early reception". In: I. Grattan-Guinness (ed.), *History in mathematics education. Proceedings of a workshop held at the University of Toronto*, Paris (Belin), 152-168.
- Rostand, F. 1960. *Souci d'exactitude et scrupules des mathématiciens*, Paris (Vrin).
- Rostand, F. 1962. *Sur la clarté des démonstrations mathématiques*, Paris (Vrin).
- Tymoczko, T. 1985. *New directions in the philosophy of mathematics*, Basel (Birkhäuser).
- Vercelloni, L. 1988. *Filosofia delle strutture*, Florence (La Nuova Italia).
- Viète, F. 1591. *In artem analyticam [sic] isagoge ...*, Tours. Also in *Opera mathematica* (ed. F. Schooten), 1646, Leiden, 1-12. [English translation: J. Klein, *Greek mathematical thought and the origin of algebra*, 1968, Cambridge, Mass. (MIT Press), 315-353.]

- West, B. J. 1985. *An essay on the importance of being nonlinear*, Berlin (Springer: Lecture Notes in Biomathematics, no. 62).
- Wigner, E. 1960. "The unreasonable effectiveness of mathematics in the natural sciences", *Communications in pure and applied mathematics*, 13, 1-14.
- Wise, M. N. 1981. "The flow analogy in electricity and magnetism. Part I. William Thomson's reformulation of action at a distance", *Archive for history of exact sciences*, 25, 19-70.
- Zellweger, S. 1982. "Sign-creation and man-sign engineering", *Semiotica*, 38, 17-54.



## **Dimensions of Applicability**



# Applying Mathematics and the Indispensability Argument

MICHAEL D. RESNIK (North Carolina)

## 1. Introduction

This paper is about applying mathematics in science and practical life. Let me begin by explaining why the topic concerns me.

According to Quine and Putnam, appealing to mathematical objects and mathematical truths figures indispensably in using mathematics in science, and as a consequence, we should consider mathematical objects to be no less real than scientific ones. In his provocative book, *Science without Numbers*, Hartry Field challenged this influential argument by offering a novel account of how mathematics might be applied even if its objects do not exist and its principles are false. In my opinion, Field's program has confronted so many technical and philosophical difficulties that it no longer constitutes a viable challenge to mathematical realism. Still, the question of whether there are philosophically attractive and technically available ways around the Quine-Putnam indispensability argument still haunts me. Perhaps, the Quine-Putnam account of how we apply mathematics is importantly inaccurate, or perhaps it is too simple. Nancy Cartwright's *How the Laws of Physics Lie* reinforced my first worry, while Henry Kyburg's *Theory and Measurement* underscored the second one.

The diversity of the Quine-Putnam, Field, Cartwright and Kyburg accounts also made me wonder whether one might apply mathematics in diverse ways — some favoring the realist viewpoint, others detracting from it. I think it is likely that this is so. But I will not try to show this here; for my knowledge of science and engineering is not up to the task, and, what is more crucial, neither realists nor anti-realists need be concerned with canvassing all the possible ways we actually apply mathematics to demonstrate their respective cases. Realists, who want to use the indispensability argument, need only show that those parts of mathematics they accept as real are



indispensable in some applications or other; while anti-realists must show only that we could achieve the same non-mathematical results without countenancing mathematical objects and truths. My plan instead is to examine the approaches that appear to undercut the Quine-Putnam argument and to argue that in so far as each succeeds it still presupposes substantial commitments to mathematical objects and principles.

## *2. Two Examples*

In order to fix our ideas and to illustrate points that will arise later, let me lay out two simple examples of mathematical applications. The first is a hum-drum, barnyard use of arithmetic and counting theory; the second is a derivation of the important, but mathematically elementary, Hardy-Weinberg law for Mendelian populations.

Turning to the barnyard, suppose that I tell you that on our farm we have three cats, four dogs and four horses, and you remark, " $3 + 4 + 4$ , that's 11, and  $3 < 4$ , so you have at least 11 animals and fewer cats than dogs". You have explicitly appealed to one arithmetical equality and one inequality; and, amongst other things, you have implicitly assumed counting principles, linking arithmetic to your numerical judgments. For example, it is reasonable to take you as assuming that if  $n < m$  and the  $F$  number  $n$  and the  $G$  number  $m$ , then there are fewer  $F$  than  $G$ . Thus a reconstructed version of your inference uses several mathematical principles in addition to the facts with which you began.<sup>1</sup>

Examples such as this one have inspired those who think mathematics is at best a theoretically dispensable, short-cut method for reasoning about scientific and practical matters. This is because it is well known that we can represent numerical quantification in first-order logic with identity and show that the premises "We have three cats, four dogs and four horses" and "Each of our animals is exactly one of a cat, dog or horse" logically imply the conclusion "We have at least eleven animals".

The example also illustrates the sort of difficulties which arise in attempting to expunge all mathematics from science; for the numerically comparative

---

<sup>1</sup> Notice that these counting principles can be derived as theorem schemata within the pure first-order theory of counting or as universal quantifications within its second-order counterpart. However, applying them requires supplanting their schematic letters (variables) with barnyard predicates (sets).

quantifier "there are fewer \_\_\_ than \_\_\_" and inferences turning on it exceed the bounds of first-order logic. In the eyes of many, using it would bring in genuine mathematics. One can get round this, by taking "fewer than" as a primitive logical operator or by defining it in second-order logic. Either way yields your second conclusion as a logical implication of non-mathematical premises, but does so only at the cost of increasing one's philosophical commitments.<sup>2</sup>

Turning to the second example, population genetics is concerned with measuring and predicting the distribution of genes. The Hardy-Weinberg law, often compared with Newton's first law of motion, is a fundamental, equilibrium principle of the field. Let me use a nice passage of Theodosius Dobzhansky's to introduce it.

Suppose that two strains of a sexual and cross-fertilizing species are introduced into a previous unoccupied territory, in which they are equally adapted to live. Suppose further that they differ in a single gene, one strain being  $AA$  the other being  $aa$ , interbreed at random, and are introduced in proportions  $p$  of  $AA$  and  $q = (1-p)$  of  $aa$  individuals. We assume that the individuals composing the population contribute equal numbers of gametes, some carrying the gene  $A$  and others its allele  $a$ , to the gene pool of this population. What, then, will be the frequencies of  $A$  and  $a$  in the gene pool, and what will be the proportions of the homozygotes,  $AA$  and  $aa$ , and of the heterozygotes,  $Aa$  in this Mendelian population in the next generation and the following ones?<sup>3</sup>

The Hardy-Weinberg law answers this question by stating that if a given breeding population is not subjected to evolutionary forces, such as gene mutation or selection, and mating is random, then the allelic frequencies (here  $A$  and  $a$ ) will remain constant from generation to generation and the genotypic frequencies (here  $AA$ ,  $Aa$  and  $aa$ ) will not vary after the first generation.

The law applies to genes having any finite number of alleles, but here is a proof for the two allele case: Suppose that in the first generation the frequencies of the alleles  $A$  and  $a$  are  $p$  and  $q$ , where  $p+q=1$ . We can now calculate second-generation genotypic frequencies using the probability calculus. Randomness assures us that the probability of a parent carrying an allele is just its frequency in the parent's generation and that the genetic contribution of

<sup>2</sup> For a case in favor of the first way see Field, *Science without Numbers*, Princeton: Princeton UP, 1980, pp. 93–95.

<sup>3</sup> Theodosius Dobzhansky, *Genetics of the Evolutionary Process*, New York: Columbia UP, 1970, p. 99.

one parent to a zygote is independent of the other. Thus the frequency of  $AA$  is given by the product of the probabilities of each parent contributing an  $A$ , which is just  $p^2$ . Similar calculations show that the frequency of  $Aa$  is  $2pq$  and that of  $aa$  is  $q^2$ . Now the second generation frequency of the allele  $A$  is just that of  $AA$  plus  $1/2$  that of  $Aa$ , that is,  $p^2 + pq$ . But this is just  $p$ , since  $p+q=1$ .<sup>4</sup> Similarly, the frequency of the allele  $a$  in the second generation is again  $q$ . Since under the hypothesis of the theorem, the allelic frequencies determine the genotypic frequencies of the subsequent generations, the genotypic frequencies of all generations after the first will be identical.

### 3. *The Quine-Putnam Account*

Before looking at anti-realist views of applying mathematics we ought to have a clear view of the Quine-Putnam account itself. On this view mathematics is applied to a particular subject by enriching our descriptions and extending our inferences. First, we increase the expressive power of the language of the target subject by adding mathematical terms to its vocabulary and mathematical objects to its range of variables. This will allow us to introduce such concepts as acceleration and state vector into physics, random mating and allelic frequency into genetics, expected utility and welfare function into economics. Second, we use mathematical laws together with non-mathematical premises to derive non-mathematical conclusions.

This account certainly accords well with a face-value reading of our two examples. By expanding our vocabulary with mathematical terms we succeed in counting my farm animals and describing the distribution of alleles in Mendelian populations. By arguing from premises drawn from arithmetic and probability theory, we manage to arrive at the Hardy-Weinberg law and mundane comparisons between my dogs and cats.

Before leaving the Quine-Putnam account I should emphasize that an anti-realist could grant that we cannot do science and engineering without using mathematical terms, variables and even existential laws, but avoid the Quine's and Putnam's realist conclusions by giving an anti-realist account of mathematical language. Only if we take mathematical names and quantifiers as having

---

<sup>4</sup> Let the total number of gametes be  $Tg$ . Define the frequency of  $AA$  (respectively,  $Aa$ ,  $aa$ ) as the ratio of  $AA$ -gametes ( $Aa$ -,  $aa$ -) to  $Tg$ . Let the total number of alleles occurring in gametes be  $Ta$ , and define the frequency of  $A$  as the ratio of occurrences of  $A$  in gametes to  $Ta$ . Note that  $Ta=2Tg$ . Thus the frequency of  $A$  is  $2\#(AA)/Ta + \#(Aa)/Ta = \#(AA)/Tg + \#(Aa)/2Tg = \text{Frequency}(AA) + 1/2 \text{ Frequency}(Aa)$ .

their standard interpretations and only if we take the mathematical premises as literally true do the realist conclusions follow. Philip Kitcher and Charles Chihara seem to accept the above account of applied mathematics but they avoid its realist implications by offering anti-realist interpretations of the language of mathematics and its subject matter.<sup>5</sup> I will not deal with this sort of response to Quine-Putnam in this paper.

#### 4. Anti-realist Accounts: Field's Structuralism

Anti-realists need a complete account of how to purge commitments to mathematical entities from mathematical applications. The account need not provide a uniform method applicable to every branch of mathematics or science – it might be a disjunctive, hodgepodge – but it must be complete. Anything short of this will fail to dismiss the prospect of mathematics being indispensable in those applications the account neglects. Of course, one tends to look first for homogeneous accounts, since these are easier to present and study. I will restrict my attention to such accounts here, but my criticisms of them would probably apply to mixtures of these views as well.

The first approach I will consider is the structural one inspired by Hilbert's *Foundations of Geometry* and the work on measurement theory codified by Krantz, Luce, Suppes and Tversky. The leading idea here is that applying a branch of mathematics to a given target domain depends upon the target domain having a structure homomorphic to a structure treated by the mathematics being applied.

On this approach, using the real numbers to measure certain bodily lengths on a ratio scale depends upon these bodies standing in a (so-called empirical) relation of comparative length that is a weak order, monotonic under bodily juxtaposition, and so on. Furthermore, where such (empirical) structures are absent, so is the corresponding possibility for measurement. Thus we cannot treat putting soap bubbles together as an operation supporting an additive measure, simply because combining two soap bubbles is unlikely to yield a third one at all, much less one whose size is in any reasonable sense the sum of the first two.

In practice, we make no clear distinction between "empirical" structures and the mathematical structures in which we embed them. No practical purpose is served by distinguishing between, say, the numerical *greater than* relation

---

<sup>5</sup> See P. Kitcher, *The Nature of Mathematical Knowledge*, Oxford, Oxford UP, 1983 and C. Chihara, *Constructibility and Mathematical Existence*, Oxford, Oxford UP, 1990.

holding between numerical values of a length function and the empirical *longer than* relation. But for theoretical or foundational purposes it may be worth attempting to characterize a target structure in terms that do not make use of coordinate systems or numerical scales. Since such characterizations contrast with their numerical counterparts as do synthetic and analytic versions of geometry, the former are commonly referred to as synthetic and intrinsic, the latter as analytic and extrinsic.<sup>6</sup> Besides geometry, synthetic characterizations can be given for certain space-time theories, measurement theory and utility theory. In each case, representation and uniqueness theorems connect synthetically specified structures and their analytic images.

(By the way, such theorems are not necessary for applying mathematics. The target domain must carry the appropriate structure, to be sure, but the structure need not be characterizable in synthetic or extrinsic terms. Furthermore, the structuralist approach as found in Krantz et. al. is compatible with the Quine-Putnam account presented above. In itself it does not undercut the indispensability argument or favor mathematical anti-realism).

Let us work through this approach using our two examples. We ordinarily take counting to be a matter of assigning numbers to finite sets. Thus the only structure presupposed by the barnyard case is that our animals, dogs, cats and horses form four finite classes. Synthetically, we can even do without the classes, as we have already seen. For we can use numerical quantifiers to formulate synthetic counting statements without referring to classes or numbers, so long as our predicates have fixed, finite extensions.

Plainly the Hardy-Weinberg case is more complicated. We may assume that we are dealing with finite biological populations; hence the frequency of the  $R$  in  $S$  is  $m/n$  just in case  $n$  times the number of  $R$  equals  $m$  times the number of  $S$ . Thus measuring frequencies amounts to counting sub-populations and comparing the results, which requires no more supporting structure than the barnyard case. But the Hardy-Weinberg law also speaks of random matings. Construing these matings as events supporting a probability measure would entail introducing a fairly complex, probabilistic event

---

<sup>6</sup> An extrinsic characterization of a structure refers to a representation of the structure in some other structure. E.g., extrinsic characterizations of spatial structures refer to co-ordinate systems. Intrinsic characterizations refer only to elements of the structure or constructions built from them. Analytic characterizations refer to numbers, functions and sets; synthetic characterizations contain no such references. Field's characterization of Newtonian spacetime is both synthetic and intrinsic. On the other hand, a characterization in terms of tensors could be intrinsic yet analytic.

structure.<sup>7</sup> However, the law only appeals to randomness to insure that a) alleles are distributed among gametes according to the frequency among the parental generation and b) the allelic contribution of one parent is independent of the others. Thus we can formulate the law and its derivation entirely in terms of frequencies.

Can we get a synthetic version of the law? Well, that depends. Now that we have reduced the law to one about frequencies, the question turns on whether we can formulate a synthetic versions of frequency statements. Consider the statement "the frequency of  $R$  among  $S$  is  $1/5$ ". This is true just in case there are 4 times as many  $S$  that are not  $R$  as there are  $S$  that are  $R$ . So by adding the primitive comparative quantifier "there are exactly 4 times as many \_\_\_\_ as \_\_\_\_" we could paraphrase the statement in arguably synthetic terms. Having taken this step, we might render "the frequency of  $R$  among  $S$  is  $3/5$ " as "there are exactly 3 times as many  $R$  that are  $S$  as one half the  $R$  that are not  $S$ ". Of course, this might require a different primitive quantifier for each rational number – perhaps even for each fraction, and we would still lack variables ranging over frequencies.

If this does not suit you, you might want to try adding "the ratio of \_\_\_\_ to \_\_\_\_ is the same as that of \_\_\_\_ to \_\_\_\_" as a primitive quantifier. Having done that, introduce the predicates, " $0x$ ", " $1x$ ", " $2x$ ", etc. defined by

$$0x \longleftrightarrow \neg(x=x); 1x \longleftrightarrow (x=a); 2x \longleftrightarrow (x=a \vee x=b), \dots,$$

where  $a, b, c$ , etc. are arbitrarily selected non-mathematical individuals. Then you could construe, say, "the frequency of  $R$  that are  $S$  is  $2/3$ " as "the ratio of  $R$  and  $S$  to  $R$  and non $S$  is the same as the  $x$  that are  $2x$  to the  $x$  that are  $3x$ ".

That is enough, I think, for it to be clear how difficult it can be to construct elegant and philosophically plausible synthetic replacements for scientific theories developed within standard, analytic mathematical frameworks. This brings us to Hartry Field, whom we should credit for making more progress with physics than our previous examples might suggest.

Field hoped to expunge mathematics from science by replacing analytic characterizations of empirical structures by synthetic ones. By maintaining that synthetic formulations contain no mathematical vocabulary, Field claimed he could refute the first part of Quine-Putnam – at least in principle, science

---

<sup>7</sup> We would probably need something like what Krantz et al. call an Archimedean structure of qualitative probability. See D. Krantz, R. D. Luce, P. Suppes, and A. Tversky, *Foundations of Measurement*, New York: Academic Press, 1971.

can be expressed without using mathematics.<sup>8</sup> As to the use of mathematics in scientific reasoning, Field planned to cover this by appealing to representation theorems. For such theorems enable us to re-describe a synthetically presented structure in analytic terms by referring to its image in some standard mathematical structure representing it. We can then use ordinary mathematics to derive properties of this representing structure, and know that, under the representing homomorphism, they transfer to the target domain in the guise of synthetic descriptions. Thus the mathematics imputes no synthetic properties to the target structure, which are not already logical consequences of its synthetic description. This allows Field to relegate the second Quine-Putnam use of mathematics to the role of a theoretically dispensable short-cut.<sup>9</sup>

Field's work is probably the most admired as well as the most carefully criticized piece of philosophy of mathematics to appear in the last 20 years. I will not attempt to review here the many problems it encountered nor Field's ingenious attempts to solve them. I will restrict myself instead to considering its success as a nominalist account of applying mathematics.

First, from the outset Field's claim that synthetic formulations are devoid of mathematics has been highly controversial. His synthetic theories are space-time theories with variables ranging over points and regions. Not only are points and regions widely regarded as mathematical entities, they have no obvious "physical" characteristics. Thus there is no convincing epistemic or ontic distinction between them and analytic entities such as numbers and sets. The difficulty only increases when we turn to synthetic formulations of theories, such as utility theoretic economics, which make essential use of continuous probability distributions or their equivalents.<sup>10</sup>

---

<sup>8</sup> Amongst other things, Field adds certain philosophical assumptions to the account of Krantz *et.al.*

<sup>9</sup> Field's program encountered a serious technical impediment at just this last step. The penultimate sentence is generally true, only if the underlying logic of the synthetic theories is at least second-order and logical consequence is relativized to standard models. First-order synthetic theories may fail to pick out a sufficiently narrow class of structures to support a representation theorem. Proof theoretic logical consequence is subject to Gödel incompleteness results. See S. Shapiro, "Conservativeness and Completeness", *Journal of Philosophy* 81 (1983), 521-531.

<sup>10</sup> The unsavory tricks I entertained in the Hardy-Weinberg case will not work here. By the way, the experts think that it is far from obvious that we can get "synthetic" descriptions of the structures used in quantum mechanics or general relativity. See G. Hellman, *Mathematics without Numbers*, Oxford: Oxford UP, 1989, p. 140. For further discussion, see my "Between Mathematics and Physics" in *PSA* 1990 vol.2, pp. 369-378.

Since any program for expunging mathematics from science is predicated upon our having a clear distinction between these subjects, it is ironic that in the course of carrying out his program Field has inadvertently undermined our confidence in such a distinction. As I see it, Field's account is a description of how so-called pure mathematics might be applied to so-called mathematical physics, economics, biology, etc. In principle, it is no different from applying one branch of pure mathematics within another – no different, for example, than the use of number theory or set theory in mathematical logic.

The second problem with Field's account as a theory of applying mathematics is related to the first: Synthetic structural descriptions of the Field type, like uncontroversially mathematical ones, remain disconnected from measurement and observation; so we still lack a full account of how mathematics is applied at the most empirical levels of science. (As Pieranna Garavaso has noted, this problem affects not only Field's program but also realist accounts based upon the structural approach of Krantz *et. al*). To appreciate the difficulty, consider how the Hardy-Weinberg law is actually used. The precise gene frequencies observed in the first generation cannot change in subsequent generations without contravening one of the hypotheses of the law. Yet measuring gene frequencies invariably reveals changes. Must this mean that biologists hardly ever observe populations in Hardy-Weinberg equilibrium? Field as well as Krantz *et. al.* simply do not address the question of how we are to deal with this kind of practical difficulty when applying mathematics. On the other hand, we can still countenance true Hardy-Weinberg equilibria, if we follow biological practice and use statistics to put some slack in the link between theoretically established equilibrium frequencies and measured ones. Then we reject the hypothesis of Hardy-Weinberg equilibrium only if "statistically significant" observations contravene it.<sup>11</sup>

### 5. Anti-realist Accounts: Kyburg's Statistical Approach

Henry Kyburg's *Theory and Measurement* is a very careful attempt to deal with the sort of problem we have just observed with respect to applying the Hardy-Weinberg law. Kyburg explicitly rejects the claim of Krantz *et. al.* (and by implication Field's) that the synthetic theories they introduce and the structures they posit are empirical. According to Kyburg, these theories have no direct observational basis. They are a priori theories to be justified by using

---

<sup>11</sup> For general help with the Hardy-Weinberg law and this point, in particular, I am indebted to my son, David B. Resnik.



a statistical error theory to show that accepting them will increase our ability to make observational predictions without significantly increasing our error rate.

Kyburg describes two ways in which we apply mathematics. First, in order to systematize some observational data and increase our predictive powers, we might construct a quantitative theory positing a quantitative structure of theoretical entities, such as lengths, and unobservable properties and relations, such as, *being a rigid body* or *being exactly as long as*. The process is like curve fitting, and in the interests of getting a useful system we may reject some of the observational data. We also stipulate true values of the quantity being measured, trying to pick values which minimize the variance in our observations as well as the rate with which we must reject observational reports. On this view, measurement by itself does not establish the true values of quantities – they are not automatically identified with means, for instance – rather true values are fixed in the process of theory construction. Moreover, we do not predict that we will observe the true value, but rather that there is a good probability that we will observe a value within a certain range of the true value.<sup>12</sup>

The second way of applying mathematics Kyburg describes begins with an extant mathematical theory with an ontology of supposed ideal entities, such as points or frictionless planes. Instead of taking the usual route and saying that certain target physical entities approximate these ideals, Kyburg suggests that we can attribute ideal properties to the physical entities themselves and treat observed deviations from these ideals as errors of measurement. As an illustration, consider the use of Euclidean geometry in carpentry. If we suppose that pencil dots are points, struck chalk lines are lines, angles subtended with a square are right angles, then within the range of acceptable carpentry errors, two points do determine a unique line, the angles of a triangle do add to 180 degrees, etc. Now, as Kyburg admits, this account faces a number of problems, and I do not know how deeply he is committed to it.<sup>13</sup> I include it because it contrasts nicely with the Cartwright view I discuss below.

---

<sup>12</sup> I have described only Kyburg's account of direct measurement (e.g., measuring lengths by comparing bodies with a standard unit body). He presents similar accounts of indirect measurement (e.g., measuring temperature by measuring the length of a mercury column) and systematic measurement (e.g., measuring voltage by measuring pointer angles on a meter constructed on the basis of electronic theory). See H. Kyburg, *Theory and Measurement*, Cambridge: Cambridge UP, 1984.

<sup>13</sup> See Kyburg, pp. 148–149. One obvious problem is that non-Euclidean geometries also fit the range of acceptable error.

Kyburg's anti-realism takes the form of a proposal for replacing variables in science that range over numbers by variables ranging over magnitudes, such as lengths, velocities, and temperatures. By construing magnitudes along the lines he recommends, we can even avoid commitments to uncountable infinities of lengths, velocities, temperatures, and other magnitudes while still closing magnitudes under the usual arithmetical operations. The trade-off is that some magnitudes, such as, two feet raised to the billionth power, may be empty. But since Kyburg does not seek representation theorems, this does not generate technical problems.

Kyburg's proposal might eliminate numbers from science, but in its present form it is hard to see how it can eliminate all mathematical entities from science. First, Kyburg uses both natural numbers and sets in setting up his magnitude structures. For example, his Archimedian axiom for length, in effect, states that if  $x$  is longer than  $y$  then some juxtaposition of *finitely many* bodies of  $y$ 's length is longer than  $x$ , and he defines irrational multiples of lengths in terms of least upper bounds of sets of rational lengths.<sup>14</sup>

Second, even if these uses of mathematical objects could be eliminated from his magnitude theories using Field-style reformulations, Kyburg would still be left with the mathematics required for the statistics he uses. As it stands this includes real-valued dispersion measures and distribution functions. Given the problems we observed in the Hardy-Weinberg case with nominalizing frequencies, the prospects for Kyburg shedding this mathematical residue are poor.

### 6. Anti-realist Accounts: Cartwright's Analogies

The final account I want to discuss I have extrapolated from Nancy Cartwright's *How the Laws of Physics Lie*. Cartwright believes that physical reality comes in messy clumps rather than in neat structures. I am certain that she would reject Field-style accounts of very theoretical physics, since she holds that its fundamental laws are fictitious. It is possible that she might endorse Kyburg's account as a fuller version of her own view. But since she does not address these issues directly, I am going to extrapolate from her view of how theoretical physics is applied to a view of how mathematics might be applied.

Cartwright thinks that we use the principles of theoretical physics as guides for constructing models of physical situations. Since these models do

---

<sup>14</sup> He also uses the least upper bound theorem in a proof on p. 86.

not depict anything real, we cannot reason deductively from the structure of the model to the structure of the real situation. Nor does Cartwright have Field's representation theorem account in mind. Instead she speaks of the models as simulacra and even of "physics as theater", suggesting to me that the reasoning in question is supposed to be analogical. Unfortunately, I cannot be sure from reading the text whether this is in fact her position.<sup>15</sup> But let us take it as hers, and see how it might run.

We want to treat mathematical models as analogs of physical situations. Yet we cannot think of them along the lines of scale models of ships or buildings or wind tunnel models, because these use one real thing to learn about another. On the Cartwright view, mathematical models are totally unreal. Now we can often appeal to totally unreal pieces of fiction to describe real situations. For example, I might refer to a TV comedy to describe politics in my state or to Don Quixote to tell you how a friend leads his life. Or I might warn you about certain business man by saying that he has a "reverse Midas touch". In each case, the truth of the story and the reality of its characters is irrelevant; to draw your conclusions all you need know are the relevant details of the story. The same idea applies to mathematical models. So long as we reason from them only analogically, we need not suppose that the mathematical objects they refer to exist or that the claims they make are true. Just as in applying fiction we only need *truth in the story*, here we only need *truth in the model*.

Consider how we might apply this view to the Hardy-Weinberg case. We could take all the problematic ideas – randomness, for instance – as mathematical constructions, and say that the law tells us that populations which are similar to the fictional ones in which matings are random will be in *something like* genetic equilibrium, provided that *something like* an absence of evolutionary forces obtains. Of course, now we must face the problem of specifying the real-life analogs of randomness, genetic equilibrium and the like without committing ourselves to mathematical objects and truths. Perhaps, this could be done operationally or in formalist terms. For example, instead of

---

<sup>15</sup> See N. Cartwright, *How the Laws of Physics Lie*, Oxford: Oxford UP, 1983, pp. 139–162. Cartwright's latest book hints at a view of applying mathematics similar to one briefly sketched by Saunders Mac Lane. On this view, applying mathematics in science is in part a matter of deducing theorems and calculating values and in part a matter of artfully and unrigorously altering formulas to fit the empirical situations. See N. Cartwright, *Nature's Capacities and Their Measurement*, Oxford: Oxford UP, 1989, pp. 212–230 and S. Mac Lane, *Mathematics: Form and Function*, New York: Springer-Verlag, 1986, p. 426.

speaking of frequencies, we might speak of the numerical symbols we write down when we count traits in a population. Instead of speaking of random breeding, we might point to the absence of certain conditions causing selective mating.

I am not sure that these suggestions can be carried out in the Hardy-Weinberg case, and I am even less confident about extending the analogical account to more recondite applications of mathematics, such as predicting precise masses of new particles.

I am also impressed by how much less precise our factual descriptions must be if we take this view seriously. Analogical descriptions have their place. I was once told that Australia is something like a less cultured version of the United States – a helpful description at the time. But in science we prize precise, quantitative descriptions and predictions over rough analogical ones.

But let us set this all to the side. Would the analogical approach succeed in ridding science of its mathematical commitments? I think not, for this reason: Although a piece of mathematics or fiction need not be true to serve as an analog, it must be consistent. Otherwise, anything will be true in it, and we could conclude anything we like by using it as an analog. So consistency is a necessity. But one lesson we have learned from mathematical logic is that we cannot establish the consistency of a mathematical theory – either semantically or proof theoretically – without assuming the existence of some objects and the truth of some laws governing them. Relative consistency proofs must bottom out in something taken as unconditionally true, even if it is only the principles of some proof theory. To be sure, the Löwenheim-Skolem theorem tells us that if we restrict ourselves to first-order theories, then we can restrict our existential commitments to a countable domain of individuals. Yet I find this of little comfort, since we must increase our structural commitments to prove the consistency of increasingly powerful mathematical systems. At the elementary levels, we can make do with quasi-mathematical objects. We might, for instance, establish the consistency of number theory using a model in discrete space-time. Yet ultimately we have no place to turn for the structures we require except to purely mathematical ones.<sup>16</sup>

---

<sup>16</sup> Just how compelling this point is depends upon how much mathematics theoretical physics requires. After discussing the issue at length, Hellman concludes that some highly theoretical pieces of physics exceed the mathematics that can be coded within second-order analysis. See Hellman, pp. 104–117.

### 7. *A Way Out? The New Modalism*

In *Science without Numbers* Hartry Field observed that we can apply analytic mathematics to a target synthetic theory without committing ourselves to mathematical objects or truths, provided that the former is conservative over the latter. By this he meant that the former implies nothing in the target language not already implied by the latter's axioms. However, he came to realize that even asserting that one theory is conservative over another, much less proving it, exceeds the limits of his nominalist framework by referring to mathematical objects, such as models or formal derivations. To get around this, Field first reformulated conservativeness as a generalized form of consistency and then proposed identifying consistency with logical possibility. (More recently, Geoffrey Hellman has taken a similar course in developing a "modal-structuralist" interpretation of mathematics). The idea here is that instead of saying that some (finite) set of axioms  $A_1, A_2, \dots, A_k$  is consistent, we say that it is logically possible that  $A_1 \& A_2 \& \dots \& A_k$ .<sup>17</sup>

This modal move may rescue the analogical approach to applied mathematics. For, according to Field, if we augment mathematics with modal logic and principles relating possibility and consistency, we can convert mathematical consistency proofs into possibility proofs without ever committing ourselves to mathematical objects.<sup>18</sup>

(Field's own approach to applied mathematics is not analogical, as we have already seen. Nor is Hellman's. The latter translates analytically formulated claims into modally formulated counterfactuals, paraphrasing, for example, "The frequency of the *A* allele is .67" as, roughly, "If there were rational numbers, then the frequency of the *A* allele would be .67". As Hellman notes, making sense of such counterfactuals encounters philosophical difficulties, but I do not have the space to explore them here).

Now Field believes that taking the modal approach yields an epistemologi-

---

<sup>17</sup> Field extends to this axiom schemata by using substitutional quantification. See H. Field, "Is Mathematical Knowledge Just Logical Knowledge?", *Philosophical Review* 93 (1984), 502–552. Hellman uses second-order versions of number theory, analysis and set theory to achieve finite axiomatizability. For further discussion of Field's views see my "How Nominalist is Hartry Field's Nominalism?", *Philosophical Studies* 47 (1985), 163–181 and "Ontology and Logic: Remarks on Hartry Field's Anti-platonist Philosophy of Mathematics", *History and Philosophy of Logic* 6 (1985), 191–209. See also C. Chihara, *Constructibility and Mathematical Existence*, Oxford: Oxford UP, 1990.

<sup>18</sup> Chihara raises a number of serious objections to this claim. See pp. 261–272.

cal gain over positing mathematical entities outright. I remain unconvinced. Of course, in most cases it is easier to show that something is possible (or consistent) than to show that it is actual (or true). Thus one might suppose that it is easier to explain how we could know that mathematical entities are possible than it is to explain how we could know that they exist.

In particular, one might suppose that we could explain our knowledge of the possibility of the natural number sequence in terms of our knowledge of the possibility of potential infinities, and that we could explain the latter in terms of our knowledge of natural processes. Indeed, Hellman writes, "Now, in fact, there may be no reason to accept [the assertion that an infinite sequence of physical objects exists], but there is every reason to accept that, logically, it might be true ..."<sup>19</sup>

But if we set aside our mathematical knowledge, where could knowledge of the possibility of potential infinities come from? Observation, biology and engineering do not tell us that natural biological or mechanical sequences can always be prolonged by one more step. Rather they tell us that prolonging a relative short sequence may be quite different from prolonging a very long one, for eventually the process of generating additional steps will exhaust the material needed or the mechanism involved.

But these limitations are biological, practical or technical. If we remove such limitations, wouldn't it be physically possible to prolong certain natural sequences? That depends upon the process involved in the prolonging. We cannot increase the velocity of a process indefinitely, nor divide matter into ever smaller parts, nor use evermore matter or energy. This rules out the physical possibility of some of the more obvious sources of potential infinities, such as, sequences of marks. Moreover, if some of the more speculative cosmologies are correct, even physical space-time is bounded. This would imply that no physical process can be continued indefinitely. Whether or not this is true, these considerations argue against using either untutored physical intuition or physical theory to ground our knowledge of the possibility of potential infinities.<sup>20</sup>

(By the way, I am not denying that our belief in the possibility of potentially infinities probably originated in our untutored physical intuitions. My point is just that these can no longer justify this belief).

---

<sup>19</sup> Hellman, p. 30.

<sup>20</sup> The possibility of geometrical ideals is even more remote from experience. It is not logically or mathematically possible to start with something spatially extended and reduce it by stepwise division to something without extension. Thus we cannot think of points as constructed by making smaller and smaller dots.

Fans of Field will be quick to reply that the relevant possibility is not physical but logical. Yes, but how do we know that potential infinities are logically possible? Because no contradiction follows from the supposition that they exist, I presume. And how do we know this? Well, we can rest with logical intuitions or we can turn to mathematical models. Historically we have taken the latter course and have appealed to mathematical objects to clarify intuitive notions of possibility. Thus possible paths and shapes gave way to curves in space, possible sizes, weights, and temperatures to abstract magnitudes and universal physical possibilities to distributions of matter in space-time. Even in logic we have turned from untutored notions of possibility to mathematical notions, replacing the idea of logical possibility with that of implying no contradiction, and explicating that in turn in terms of proof theoretically defined deductions or set theoretically defined interpretations. Moving to the more abstract realm of mathematics allows us to put our ideas in their simplest and most uncluttered forms, giving us thereby the best chance of determining their consistency. It is easier to determine whether the idea of the natural number sequence qua bare structure of the (potential) infinite harbors a contradiction than it is to determine whether the idea of some physical infinity does, simply because in the first case we need not worry about the effect of extra physical baggage. Seen in this light, paraphrasing talk of consistency by possibility retrogrades, replacing the clear by the obscure and the methodologically advanced by the intuitive.

Someone might object at this point as follows. To answer your modalist, you must establish that it is no harder for us to account for our knowledge that mathematical entities *exist* than it is for us to account for knowledge that certain physical ideas are *consistent*. Yet so far, the most you have established is that we can more easily account for our knowledge of the consistency of certain mathematical ideas than for our knowledge of the consistency of related physical ideas. Granted, but remember that once we replace the unexplicated idea of logical possibility by that of consistency, then committing ourselves to the *consistency* of the idea of the natural number sequence also commits us to the *existence* of equally complex mathematical structures required to explicate the notion of consistency. This is why it is essential for the modalist to abandon the possible world semantics and take modal operators as unexplicated primitives.

Turning from epistemological issues to metaphysical ones, it is difficult to assess the import of the modal twist. In general, possibility is weaker than actuality, but when it comes to mathematical objects this is no longer evident. Some philosophers believe that the same mathematical objects exist in every possible world, if they exist in any. For them, mathematical possibility

implies mathematical existence. On the other hand, Field, who rejects the possible world approach, takes himself to be concerned with the broader notion of logical possibility instead of the metaphysical notion used by possible worlds theorists.<sup>21</sup>

Even when we hold the notion of possibility fixed, we lack a single way of thinking about the possibility of mathematical entities. Some philosophers treat the issue as analogous to the possibility of unicorns or tachyons. On this view, mathematical entities are objects that could exist but do not. Some might add that they could even exist without affecting the nonmathematical universe. Other philosophers, whom I will call Aristotelean structuralists, hold that mathematics is about omega sequences, complete ordered fields, iterative hierarchies, and other structures, and that these exist if and only if they are instantiated. To such philosophers saying that the natural numbers are possible is really to say that omega sequences are possible, which, in turn, is to say that it is possible for some (presumably nonmathematical) things to form an omega sequence. (This seems to be Hellman's view). Finally, Platonic structuralists also hold that mathematics is about structures, but allow for structures to exist uninstantiated by nonmathematical entities. For these philosophers the distinction between possible, uninstantiated structures and actual, uninstantiated structures lacks content. This is how I am inclined to approach the issue of the possibility of mathematical objects.

Unfortunately, I have no simple argument to offer for my particular approach to structuralism, and the case I would offer were I to have more time is far from air-tight. So I will conclude now in the hope that I have been able to make it clear that purging mathematical objects and truths from their customary applications is no easy task.<sup>22</sup>

---

21 These considerations do not apply to Hellman's view or at least do not do so directly. Hellman does not postulate that specific mathematical objects are possible but rather that it is possible that certain structures are instantiated. In particular, he does not posit the possibility of the natural numbers but only the possibility of something being an omega sequence, and the something in question might be concrete.

22 I would like to thank Pieranna Garavaso, Keneth Reed, David Resnik and Geoffrey Sayre McCord for their help with this paper.



## Mathematical Structures and Physical Necessity

ROBERTO TORRETTI (Puerto Rico)

In the few languages I am familiar with there are standard ways of expressing the necessity of an event or a state of affairs, and also a noun – ‘necessity’, ‘necesidad’, ‘Notwendigkeit’, ‘Ανάγκη’ – for naming it. I do not know whether this is a feature shared by every human language, but it appears to be well entrenched in every language of the European tradition within which mathematical physics was born and continues to be nurtured. I take this to mean that speakers of these languages perceive or think they perceive in some events and situations a distinctive feature which requires that mode of description. There can be no question that people who grow within the said linguistic tradition do articulate their experience in such a way that it displays the appearances of something I shall call ‘perceived necessity’.

There are several kinds of perceived necessity. First of all, there is the necessity of the past. A foul in a football game might have been avoided, can still be penalized, but cannot be undone. And the same is true of course of every particular event and action, no matter how random or free. The necessity of the past was hotly debated among medieval theologians, who encountered here an unbreakable limit to their God’s omnipotence. It has also been the subject of some great poetry, as in Shakespeare’s metaphor, “All the perfumes of Arabia will not sweeten this little hand” (*Macbeth*, V.i.47), or in Pindar’s lines

τῶν δὲ πεπραγμένων  
ἐν δίκᾳ τε καὶ παρὰ δίκαν, ἀποίητον οὐδ’ ἂν  
χρόνος ὁ πάντων πατὴρ δύναιτο θέμεν ἔργων τέλος.

Of things consummated, whether just or unjust, not even Time the father of all can make the end undone.

(Ol. II 15–17)

On the other hand, physics has remained notoriously indifferent to it, due perhaps to the fact that such a drably homogeneous and utterly pervasive form of necessity makes no difference in reality; or because, as every particular attains it merely by occurring, this kind of necessity lacks the specific, i.e. restricted,

universality which, since Aristotle, natural science has considered to be its proper object. Be that as it may, mathematical physics has dealt exclusively with two other forms of perceived necessity, and I shall therefore concentrate on them.

One might refer to them informally as present necessity and future necessity, thus completing a nice match with the familiar classification of times; but such terms would be misleading. So I shall rather call them the necessity of configurations and the necessity of processes. We perceive the former, for instance, if we try to tile our kitchen floor with regular pentagons of the same size and soon realize that it is impossible.<sup>1</sup> We perceive the latter when we are knocked down and dragged by a strong wave while bathing in the sea; or when we apply an electric saw to a thin beam of wood and promptly cut in two; or, only a little less obviously, when we see the flame of a match catch a small piece of paper and turn it to ashes.

Let me emphasize that I am not speaking of things as they are in themselves – of which I confess that I have no inkling – but as they appear in the daily lives of men and women like ourselves, who speak one of the languages rich in idioms of necessity to which I referred earlier. We often employ such idioms to describe processes of the sorts I have evoked, and although our descriptions may sometimes turn out to be inappropriate, we can judge them so only by comparison with other like descriptions which provide standards for the right use of modal idioms.

David Hume maintained that, since nobody can sense a necessary connection between the successive stages of an ongoing process, the apparent necessity of some processes merely reflects the habitual expectations of the perceiver. But this is only the conclusion of a philosophical argument based on questionable premises about the nature of sense awareness. And Hume's argument collapses if in fact we sense flows and not just static qualities; if, as one might say in Newtonian jargon, I feel the rate of change of my body's momentum when I am pushed by a wave. On the other hand, Hume presumably admitted the perceived necessity of configurations under the heading of "relations of Space and Time", one of his seven sources of "philosophical relations". Evidently, mathematical physics makes no such distinction between these two forms of necessity. On the contrary: from its inception in the 17th century it has systematically sought to extend to natural

---

<sup>1</sup> After writing the above, I was pleased to learn that James Franklin ("Mathematical Necessity and Reality". *Australasian Journal of Philosophy*, 67 (1989), p. 286) uses precisely the same example to illustrate the relevance of mathematical necessity to physical reality. I thank Professor Franklin for sending me a copy of his paper.

processes the intellectual grasp of necessary configurations bequeathed by the Greeks under the name of "geometry". This shows, by the way, how deeply alien the Humean philosophy was to the program of modern physics.

Geometry's way with configurational necessity is well known. Any configuration of interest is analyzed into points, curves and surfaces residing in a background milieu or "space" and satisfying an overseeable set of conditions as to the relations they have among themselves and with other like residents of the space. The necessity of a configurational feature is *understood* if it is shown to follow from those conditions. Two ideas deserve our attention. FIRST, the set of required conditions must be *overseeable*, by which I mean that it must be either finite or recursive and that, if finite, it must be small (geometry would have little use for a structure sporting, say, one million primitive relations characterized by one trillion independent elementary conditions), and if recursive, it must be defined in a fairly simple way. The notion of overseeability must be kept vague for there is obviously no way of setting a maximum to the cardinality of overseeable finite sets or to the complexity of overseeable recursive definitions. SECONDLY, to understand the necessity of a feature perceived in a configuration is tantamount to grasping that feature as a necessary consequence of the characteristics by which that configuration is conceived. In other words, the perceived necessity of a feature becomes understood necessity as soon as it is anchored to our chosen concept of the configuration that displays that feature. (Of course, many initially unperceived necessities have also become known in this way). The important thing here is the idea of necessary consequence, or rather, that of grasping something as a necessary consequence of a set of conditions. The connection between the necessitating conditions and the necessitated feature is made manifest – or, as the saying goes, *proved* – by arguing from a suitable description of the former to a suitable description of the latter. Such arguments consist of words, accompanied by drawings or even by gestures. Because of it, understood necessity is often called *logical*, i.e., verbal, or *dialectical*, i.e., argumentative, necessity. This is not improper, but it is apt to be very misleading, given the fact that in current usage, 'logical arguments' are those that can be faithfully represented by certain computable sequences of well-formed formulae in a so-called logical calculus. Now, Gödel showed almost sixty years ago that it is hopeless to try to codify in this fashion all proofs leading from overseeable sets of conditions to their necessary consequences. Therefore, if one must cater to the traditional preference for Greek terms, one ought to say that understood necessity is *dianoetical*, not just logical; or, recalling that one of the meanings of the Greek verb *μανθάνω* was 'to understand', one should simply call it *mathematical necessity*.

At first blush, the static, enduring necessity of configurations has precious little to do with the necessity of processes. The latter is not just timebound but timebinding: it ties the future to the present and the past, it forces change and preordains its outcome, channelling the flux of events. Yet the amazing contribution of modern physics to the exploitation of natural processes by man has resulted from understanding their necessity on the analogy of geometric necessity, through the representation of such processes by mathematical structures. This was made possible by one of the most remarkable feats of human thought: the conception of time – past, present, and future – as a single, homogeneous linear continuum.

A lucid explication of what it takes to be a linear continuum was not available until the late 19th century but the science of motion established two centuries earlier by Newton clearly presupposes that time has the same structure as the trajectory of a free particle in space. Indeed, when Galileo, in the Third Day of the *Discorsi*, let a line represent the time in which a certain space is traversed by a body in uniformly accelerated motion and used the geometrical properties of that line to establish a functional relation between travelled distances and travel times he must have understood that the time can be mapped bijectively onto the line, so that each segment discernible in the latter uniquely corresponds to a distinct subinterval of the former.<sup>2</sup> A similar understanding was already implicit in the use of geometrical methods in Greek astronomy and also in the use of kinetic methods in Greek geometry itself. For instance, Archimedes studied a spiral drawn by a point receding from another point with constant speed along a straight line that rotates about the second point with constant angular velocity. The construction assigns a definite point on the plane to each instant of the motion, viz., the position reached at that instant by the moving point. Archimedes' spiral is thus defined by an injective mapping of a time interval into space. Similar parametrizations are involved in Greek models of planetary motion. The earliest and simplest were devised by Eudoxus, who associated each planet with an  $n$ -tuple of concentric spheres, such that (i) their common center is the center of the Earth; (ii) the first sphere rotates about the poles in one day; (iii) for each  $k$  greater than 1 and equal to or less than  $n$ , the  $k$ th sphere rotates with a constant speed of its own about an axis fixed on the  $(k-1)$ th sphere; (iv) the planet lies on the equator of the  $n$ -th sphere. If the number of the moving spheres and their respective axes and speeds of rotation are suitably assigned, the astronomer should be able to calculate from the planet's current position

---

2 Galileo, *Le Opere*. Nuova ristampa della Edizione Nazionale. Firenze: G. Barbera, 1964–1966, vol. VIII, pp. 208–10; cf. p. 85.

its declination and right ascension at any time. The description of a particular Eudoxian model can be read as the definition of a special kind of curve – e.g., a Venusian or a Martian trajectory – drawn by a point on any spherical surface, such as the vault of heaven or the ceiling of a planetarium. A body affixed to such a point instantiates the Eudoxian concept of a certain planet. The radius vector of the planet points in such-and-such directions at such-and-such times with the same inexorable necessity as the position vector of an Archimedean spiral reaches at prescribed times the third, the fourth,...the  $n$ th turn of the curve, or as any side of an equilateral triangle makes internal angles of  $60^\circ$  with the other two. The future path of the bright spot we are now observing is fixed and timed by what it *is* at present, if indeed it *is*, say, a venus or a mars in the sense aforesaid.

Is the obvious difference between modern physics and ancient astronomy merely a matter of complexity ... and predictive success? Or have Newton and Maxwell, Einstein and Dirac, reached for – and sometimes attained – an essentially deeper understanding of events than Eudoxus or Hipparchus? Much of the 20th century debate in the philosophy of science turns, openly or covertly, around this issue. For my own part, I am inclined to view the Eudoxian program as a paradigm of mathematical physics, *provided that it is interpreted realistically*. There are, however, some important differences between it and the generally acknowledged Newtonian paradigm. Let me mention three.

(1) Newton's physics is not a collection of (related) mathematical recipes for the prediction of regular occurrences in nature but a system of mathematical principles for natural philosophy. Its basic notions of time and space, mass and force, bound together in the Laws of Motion, are ultimately meant to account for *all* phenomena, not just for those observable in this or that part of our environment.

(2) In all theories designed after the Newtonian paradigm, the states of physical systems are described and their evolution is explained in terms of purportedly universal properties of matter. Specifically, Newton's unified theory of planetary motion and free fall subsumed these two seemingly disparate families of physical processes by postulating a mutual attraction, dependent solely on mass and distance, between all pieces of matter. Second-generation Newtonians took this to mean that each piece of matter was inherently the seat of an attractive force, acting on every other piece of matter throughout the entire universe. Subsequently, physicists postulated further universal forces to account for electric and nuclear phenomena. Current research aims at unifying all acknowledged elementary forces in a single theory.

(3) In Newtonian physics the future and past states of a physical system are

linked by differential equations to the forces actually working on it.

Since the unifying "theory of everything" could well remain an elusive dream and the very idea of a "force of nature" may prove to be disposable, the essence of mathematical physics must properly be seen to lie in feature (3), specifically, in the representation of natural processes by time-dependent functions or by sections of bundles over spacetime, which satisfy suitable differential equations. I do not have to explain here how the properties of some differential equations make them wonderfully fit for providing an intellectual grasp of necessary processes. As you all know the theorems on the existence and uniqueness of solutions ensure that, when their conditions are met, a process that can be adequately represented by an integral curve of a differential equation is indeed inexorable. I shall therefore take the mathematics for granted and make some reflections concerning its application to phenomena.

The mathematical representation of natural processes in the manner indicated affords a clear and simple answer to the philosophical question raised by Kant, presumably at the prompting of Crusius or Hume. Kant asked: "How shall I understand that because something exists something else should exist as well?".<sup>3</sup> Why, for example, should a twinkling light make its first appearance on our sky tonight because a star exploded in the Andromeda galaxy two million years ago? If the two realities under consideration can be conceived as stages in a process governed by a differential equation that determines a unique integral curve through the point representing one of those stages, then the other stage necessarily must correspond to another point on that curve and therefore – given its spacetime location – it cannot be different from what it is. Note that this answer presupposes that the distinct realities linked by a necessary connection are grasped as phases or aspects of a single reality that encompasses them, namely, a single, isolated physical system evolving in accordance with a definite law. Mathematical physics is monistic, at least within the bounds of each one of its applications. No wonder then that time and again physicists have reached for a unified theory of the whole world, in which, say, every discernible feature of phenomena is encoded by the local value of a section of a sufficiently complex fibre bundle. All such features would then be tied to one another with bonds of necessity by the differential equation defining that section. Perhaps such a theory is just a dream. But anyway inside a given application mathematical physics will not countenance the kind of pluralism one finds in fairy tales; I mean the kind of situation in

---

<sup>3</sup> "Wie soll ich es verstehen, daß, *weil etwas ist, etwas anders sei?* " (Kant, *Gesammelte Schriften*. Herausgegeben von der K. Preußischen, bzw. Deutschen Akademie der Wissenschaften. Berlin, 1902ff., vol. II, p. 202).

which, say, a physical system subject to the laws of frog physiology is abruptly replaced by another one which follows the rules of princely demeanor; or a rotting corpse suddenly feels alive and well and returns to his job; or a shining palace of marble and gold is created *ex nihilo*. Such situations cannot be judged impossible while there is no satisfactory theory of the whole world to exclude them, but they remain entirely outside the scope of physics. And, of course, if any such situation did in fact arise, it would be seen as a breakdown of the natural necessity that physics seeks to understand.

Now the very answer thus given to Kant's question of existence prompts a thorny question of knowledge. It is all very well, we might say: if the Solar System is to a certain approximation a sufficiently isolated 10-particle system governed by Newton's Law of Gravity, then, if the planets have such and such positions relative to the fixed stars tonight at 1 am GMT, they must have occupied such and such other positions exactly 10 years ago. But how can we know that the Solar System is in fact such a system? Generally speaking, how can we ascertain that some actual state of affairs we have singled out in Nature's flux should be regarded as a stage in a physical process governed by such and such differential equations? If 'to ascertain' means to establish with certainty, the answer is plainly that we cannot. The risk of error ordinarily involved in bringing a given particular under a universal concept is compounded in this case by two specific difficulties. Suppose that the information we have about the changing state of affairs under study can be satisfactorily represented by an extensible curve in a suitable differentiable manifold – be it spacetime, or some bundle over spacetime, or an abstract phase space. Obviously, such a curve can be extended in many different ways, matching apposite solutions of diverse differential equations. Hence, all the information we can gather about an evolving physical system, no matter how accurate and exhaustive it may be, will never be sufficient to establish with certainty how that system will continue to evolve in the future. Note, however, that this is so, not because – as a Humean would argue – the future might not be like the past, but because we take the successive stages of a physical system to be so tightly knit together that before the future has been realized we cannot quite know what the past was like (*viz.*, what species of structure it was an instance of). For all its inherent uncertainty, physics never faces the unfeasible task of inferring mathematical structures from raw sense data – as the Humean would have her do. She must choose between alternative conceptually articulate readings of phenomena, and this, of course, is well within the reach of educated guessing and the standard methods of statistical inference. And while the spectrum of alternatives may be infinite in principle, the choices available at any given stage of science are very few, due to our lack of imagination, the poverty of

our mathematics and the variety and complexity of phenomena. Indeed, to narrow down the viable choices is probably the best reason we have for seeking to embrace a great variety of phenomena in a single theoretical system.

I have deliberately ignored what is perhaps the most significant feature of the mathematical representation of nature, namely, that the exact concepts of mathematics must be blurred in order to fit our observations. 20 years ago Professor Günther Ludwig developed an exact theory of such blurring, based on the topological notion of uniformity.<sup>4</sup> It is a pity that philosophers of science outside Germany have not yet paid sufficient attention to his work. I cannot go into it here. Let me just recall that, due to the said blurring, the modelling of natural processes by solutions of differential equations can be used for prediction only if these solutions are stable. On the other hand, differential equations whose integral curves run into points or areas of instability, or even into singularities, are being increasingly – and very fruitfully – employed for the representation of physical processes. They are particularly appropriate for dealing with rich and intricate systems – such as one meets in meteorology or cosmology – where the clearest manifestations of necessity – e.g., the onslaught of a hurricane – baffle all attempts at prediction. Interestingly for philosophers, such mathematical models show once and for all that in so-called deductive-nomological explanations foresight and understanding do not always go hand in hand.

Allow me to finish with a half-baked suggestion. I noted at the beginning that mathematical physics wholly disregards what is perhaps the most striking distinction between items in our experience, namely, that between past and present items which are remembered or perceived, and future items which are merely anticipated. It will now be clear why this had to be so. If time is equated with a linear continuum all periods of time become essentially alike and will differ at most in length. There is, however, one physico-mathematical concept that is designed to cope with the transition between the uncertainty of the future and the necessity of the past and therefore involves the said distinction, at least implicitly. I mean the concept of probability, understood as quantified facility (in Galileo's sense)<sup>5</sup> i.e. as single-case, objective chance.

---

<sup>4</sup> G. Ludwig, *Deutung des Begriffs "physikalische Theorie" und axiomatische Grundlegung der Hilbertraumstruktur der Quantenmechanik durch Hauptsätze des Messens*. Heidelberg: Springer, 1970. (Lecture Notes on Physics, N° 4.)

<sup>5</sup> Galileo, *Le Opere*, vol. VIII, pp. 591–94. Cf. Leibniz' dictum: "Quod facile est in re, id probabile est in mente". *Sämtliche Schriften und Briefe*, herausgegeben von der Preußischen, bzw. Deutschen Akademie der Wissenschaften zu Berlin (Darmstadt–Leipzig–Berlin, 1923ff.), VI.ii, p. 492.



This concept is central to contemporary physics, inasmuch as the Quantum-theoretical representation of a physical system hinges on the notion of a probability wave that evolves with necessity according to a differential equation and yields, through its local amplitudes, the information required for computing a rich variety of interesting chance distributions. The suggestion I want to make is this: if physical probabilities quantify the comparative facility of uncertain prospects it is no wonder that a state vector encoding information on probabilities should "collapse" as soon as one of the prospects in question becomes a (henceforth necessary) fact.

# The Role of Mathematics in Physical Science

ERHARD SCHEIBE (Heidelberg)

## 1. *The Unreasonable Effectiveness of Mathematics*

Since the beginning of modern physics four hundred years ago physicists are convinced of Galileo's saying that "the book of nature is written in the language of mathematics". Concerning this it has been expressed over and over again that the usefulness of mathematics for our understanding of nature appears to be a miracle. For Kepler and Galileo the miracle was that in reading the book of nature, we are directly confronted with the thoughts of God. Wigner speaks of "the unreasonable effectiveness of mathematics in the natural sciences".<sup>1</sup> In an article bearing this title he frankly admits "that the enormous usefulness of mathematics in the natural sciences is something bordering on the mysterious and that there is no rational explanation for it". Einstein, too, speaks of the "enigma that researchers of all times have worried so much about. How is it possible that mathematics, a product of human thinking independent of any experience, so excellently fits the objects of physical reality?"<sup>2</sup> In line with these statements Steven Weinberg gives, so to speak, an empirical confirmation of the miracle by enumerating many cases where structures needed in physics had been already found and developed by mathematicians "long before any thought of physical application arose. It is positively spooky how the physicist finds the mathematician has been there before him or her".<sup>3</sup>

Is there an explanation for the miracle? Both Einstein and Weinberg attempted to give one. For Einstein it is contained in his famous words: "As

---

1 E. Wigner, "The Unreasonable Effectiveness of Mathematics in the Natural Sciences". In: E. Wigner, *Symmetries and Reflections*, Woodbridge, Conn., 1979, 222–37; the two following quotations are from pp. 223 and 229 f.

2 A. Einstein, "Geometrie und Erfahrung". In: *Mein Weltbild*, Ullstein Materialien, 1934, 21983, 119–27. The quotations are from pp. 119 f.

3 "Mathematics: The Unifying Thread in Science". *Notices of the Amer. Math. Soc.* 33 (1986) 716–33; here: pp. 725 and 727.

far as the propositions of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality". Weinberg's explanation is: "Mathematics is the science of order, so perhaps the reason the mathematician discovers kinds of order which are of importance in physics is that there are only so many kinds of order". The two explanations seem to say quite different things. In fact they belong to the same picture and complement one another. They draw on two fundamental features of modern mathematics. Einstein explains his view by the remark that it is only the recent thoroughly axiomatic orientation of mathematics that has made it perfectly clear "that through it the logically formal has been neatly separated from the material [...] content [and] that only the logically formal [...] constitutes the object of mathematics". Precisely through this separation mathematics obtains its much admired certainty. As soon as it is removed from its isolation to be applied to physical reality it becomes affected with the uncertainty of the decision *which* of the infinitely many species of structures that *could* be applied in principle has to be chosen in any given case. It is here where Weinberg can make his point. Roughly his remark is: Mathematics in its present shape offers *all* exact forms of thinking that man is capable of. By choosing one of them to solve our physical problem we are doing the only thing that can be done at all. And there is an overwhelmingly great variety. Small wonder that we reach our goal.

## 2. *An Illustration*

Does this view explain the pre-established harmony between mathematics and physical reality? Many things could be said in answer to this, and one thing, I think, has to be admitted outright. The amazing expansion of mathematics in our century, essentially, if only implicitly, used in the explanation, certainly has diminished the miracle of the applicability of mathematics to nature. In the 17th century the rejection of geometry (for whatever reason) would have meant the rejection of half of the mathematics known at the time – an irreplaceable loss. Today the abandonment of current geometry would only mean its replacement by another one – a simple transition from one species of structures to the next. This is not to say that we or our descendants, by reasons coming from physics, could never be driven to the point where we or they would be at a loss concerning present mathematics taken as a whole. It is not clear, for instance, what kind of mathematics is now used in quantum field theory. There are mathematically or physically motivated expeditions into borderlands of present mathematics such as non-standard analysis, non-cantorian set theories, many-valued logics, quantum logic and the like. These, and other conceptions still to be developed, may one day lead to a new miracle of the kind in question.

However, even the present situation leaves something to wonder about concerning what is omitted in the picture sketched by Einstein and Weinberg. And it is this something which I want to talk about in this paper. As we have seen in Einstein, he thinks that it is only the logical form through which mathematics enters a physical theory. If this were the case everything that we would want to say about physical reality could be formulated in purely physical terms aided by the usual logical operations. We could then still marvel at the fact that reality is as it is, and not otherwise. To this a Platonist could add the marvel that the physical structures found by experience one and all are isomorphic to mathematical structures. And even an instrumentalist may be surprised that all the mental pictures needed to describe nature are stored in one large unified logico-mathematical system. We could either talk in mathematical terms instead of talking in physical terms. But there would be no need to turn to mathematics, as distinct from logic, in order to formulate a physical law or to axiomatize a physical theory.

Now my question is: Is there really no need? Let me illustrate the situation by the simple example of an empirical law, e.g. a gas law. By a gas law the physicist wants to express a relation between pressure, volume and temperature of a gas that is valid for as many sorts of gases as possible. Though united in one and the same gas, the quantities in question are of entirely different kinds, and at face value it is difficult to see by what means we could formulate a relation between them. It is well known how this is actually done in physics. There is one thing that physical quantities like the ones in question have in common: Their values can be described by real numbers. With one stroke this uniformization makes possible what appeared to be impossible before: the wealth of 3-termed relations between numbers is at our disposal to formulate the law. However, for this gain we have to pay a price. Numerical relations are based on the elementary operations with numbers and eventual limiting processes, and these operations and processes have *no* physical meaning in our little theory. In other words we have attained a physical law *not* by formulating it in terms of pressure, volume, temperature *and further physical concepts*. Rather, we have embedded the given structures of three 1-dimensional value scales in a richer mathematical structure with elements having no counterparts in the physical systems to be described. And with the help of these additional elements, viz. the operations with real numbers, we succeeded in formulating relations between the physical entities from which we started. There is no question of merely adding logic to physical concepts to obtain the law.

So far I have not shown you that the method illustrated *must* be applied in order to make theorizing in physics possible. I have only reminded us of what we are *actually* doing in writing down a physical law. But even this may

already create the feeling that the method by which mathematics is used in the given example, if taken together with its empirical success, is indeed "something bordering on the mysterious". So much more when I add that I have only described what is *normally* done. It is entirely normal that in formulating physical theories we use mathematical terms for which, though these terms occur in descriptive position, no physical meaning is so much as intended. Spacetime coordinates and the associated components of forces, fields etc. are the most familiar examples. A single coordinate system has no physical meaning. But it enables us to formulate a physical law via some differential equation. Again and again physical statements are obtained by making use of numbers in this and other ways, and it is by no means obvious how the same thing could be achieved without their use. It is this feature which, to say the least, is not sufficiently articulated in the Einstein-Weinberg view.

### 3. *The Mathematical Overdetermination of Physics*

In the following I shall suggest a method by which the phenomenon of *mathematical overdetermination of physics* – as it might be called – can be investigated.<sup>4</sup> As my starting point I choose a slight modification of the common concept of a physical theory. According to the received view a physical theory is essentially a formalism provided with an interpretation. In view of the problem we have to tackle it is convenient to modify this conception by assuming not one formalism but a *pair* of such to be associated with a physical theory. In general either formalism of this pair receives a physical interpretation. Though we have thus two interpreted formalisms – *semi-theories* as I will call them – it is generally *not* the case that each of them can do the job that they do jointly. The deficiency of the primary semi-theory is the incompleteness of its physical interpretation. The deficiency of the secondary semi-theory is that it has no axioms or that its axioms are incomplete. The classical case of this kind of analysis is found in analytical geometry where the primary semi-theory is the arithmetical version of geometry and the secondary is, as far as available, the formalization of the geometrical concepts and axioms proper. For an advanced physical theory viz. quantum mechanics, the idea of such a splitting was clearly expressed (and generalized) by Hilbert, v. Neumann and Nordheim in a paper of 1928. There they said: "Certain physical requirements are imposed on the probabilities, suggested by our experience [...] and implying

---

<sup>4</sup> The present paper continues earlier considerations in my "Mathematics and Physical Axiomatization". In: *Mérites et Limites des Méthodes Logiques en Philosophie*. Ed. Fondation Singer-Polignac, Paris, 1986, 251–77.

certain relations between probabilities. Then we look for a simple analytical formalism involving quantities that satisfy just these relations [...]. The aim is to formulate the physical requirements with just sufficient completeness to define precisely the analytical formalism".<sup>5</sup>

To be more precise, let

$$(1) \quad \begin{cases} (F_1, \mathfrak{I}_1), & F_1 \equiv (S_1, \vdash_1, A_1) \\ (F, \mathfrak{I}), & F \equiv (S, \vdash, A) \end{cases}$$

be two semi-theories with formalisms  $F_1(F)$  and interpretations  $\mathfrak{I}_1(\mathfrak{I})$ . As usual each formalism  $F_1(F)$  consists of a language, represented by a set of sentences (or formulas)  $S_1(S)$ , a logic represented by its implication  $\vdash_1 (\vdash)$  and a nonlogical axiom-system  $A_1(A)$ . The language should distinguish between physical, mathematical and logical terms where the latter are neutral with respect to the difference between the former. In an obvious way this distinction leads to a division of all sentences of either formalism into physical, mathematical and mixed ones. In the interpretation  $\mathfrak{I}_1(\mathfrak{I})$  the physical terms should have physical referents. For the following considerations it does not matter whether the mathematical terms, too, refer to corresponding entities or are left without any interpretation. However, it is *not* assumed, not even for the logical case, that our division can be made on the basis of any general criteria inherent in the concept of a formalism. If there were generally acknowledged criteria drawing a sharp dividing line between mathematically and physically interpretable formalisms I would be the first to apply them. But there are none, and so our division is meant to come about with each concrete physical theory according to the interpretive intentions associated with it.

The next thing to take care of is the *relation* of the two semi-theories, welding them together to become parts of one and the same theory. The relation is given by an injective mapping

$$(2) \quad \rho : S \rightarrow S_1$$

'translating' the sentences of  $F$  into those of  $F_1$ . This translation is not meant to preserve meaning in the usual sense in every case. It does preserve physical meaning in passing from the primary to the secondary semi-theory and mathematical meaning in the opposite direction. Thus we have *conservation requirements* of the kind

- (a) If  $\alpha_1 \in \rho(S)$  is physical so is  $\rho^{-1}(\alpha_1)$ , and they have the same meaning.

<sup>5</sup> D. Hilbert et al., "Über die Grundlagen der Quantenmechanik". *Math. Annalen* 98 (1928) 1–30; here: § 1.

- (b) If  $\alpha \in S$  is mathematical so is  $\rho(\alpha)$ , and they have the same meaning – if any.

If we add to this the postulate that

- (c) All physical sentences of  $S_1$  belong to  $\rho(S)$

then the idea that in general the secondary semi-theory is more physical than the primary one is realized as far as the languages are concerned. In the ideal case the whole secondary language is physical.<sup>6</sup>

So much for the relation between the interpretations. As to the logics and axioms analysis of the physical examples show that in general the primary formalism is stronger than the secondary. This gives us

$$(3.1) \quad \begin{cases} \Gamma \vdash \alpha \Rightarrow \rho(\Gamma) \vdash \rho(\alpha) & (\Gamma \subseteq S, \alpha \in S) \\ A_1 \vdash \rho(A) \end{cases}$$

According to a widespread usage in similar cases, the embedding (2) could then be called a *representation* or *interpretation* of the secondary semi-theory in the primary one. However, in view of the actual semantical situation just indicated, this terminology is justified only in a syntactical sense. Our *main problem* is whether the inverse implications

$$(3.2) \quad \begin{cases} \rho(\Gamma) \vdash \rho(\alpha) \Rightarrow \Gamma \vdash \alpha & (\Gamma \subseteq S, \alpha \in S) \\ \rho(A) \vdash A_1 \end{cases}$$

also hold. In the terminology just introduced, the representation would then be called *conservative*. If we can find a semi-theory decomposition of a given physical theory whose secondary component is conservatively represented in the primary one, *and* if the secondary language is purely physical, then we can forget about the primary component and with it forget about an extra mathematics in the theory. If we cannot find it we have to put up with an uneliminable piece of mathematics necessary in order to formulate the theory. It is this case which, if it occurs, would appear somewhat miraculous to my mind. The presentation of such cases would be affected by the unfortunate circumstance that, while it is easy to prove the existence of a decomposition satisfying (3) once one has found one, the proof of the non-existence can be

---

<sup>6</sup> In terms first used by Ludwig the theory then has an axiomatic basis, see G. Ludwig, *Die Grundstrukturen einer physikalischen Theorie*, Berlin, 1978, § 7. However, in Ludwig's approach it is the distinction between observational and theoretical terms rather than that between physical and mathematical ones that seems to be the main concern, cf. note. 7.

very difficult, and in fact the only proofs known to me concern the logics and are founded on Gödel's incompleteness theorem.

Before looking at some details a remark is in order about how the foregoing distinction between physical and mathematical terms in a theory is related to the empiricists distinction between observational and theoretical terms. To some empiricistly minded people it may even occur that the two distinctions are identical. In response to this suggestion<sup>7</sup> it has to be admitted that there are some obvious structural similarities between the two distinctions. More important, however, is a warning. Although expressing the limits of our observational capabilities the theoretical terms of the empiricists were never exempt from receiving an interpretation by elements of *physical* reality, *besides* and *independent* of the interpretation of the observational terms. By contrast, mathematical terms have no physical interpretation besides and independent of the physical terms. As far as mathematical terms refer to anything at all their referents do not belong to the objects to which our physical theory is applied. It makes simply no sense, to give but one example, to look at a quadruple of numbers related to a spacetime point by a coordinate system as another physical entity to be treated by the theory. If we wanted to take account of the theoretical-observational dichotomy we would have to introduce a *further* division of the *physical* terms.

#### 4. Reconstructions within Set Theory

Having sketched the interplay of physical theory with mathematics in general, I can now come to some special cases. Present physical theories like classical and quantum mechanics, general relativity theory, etc., are most conveniently reconstructed on the basis of *set theory*. Accordingly, an important special case of our general scheme will be the case where both semi-theories into which a given physical theory is decomposed, are extensions of set theory. Now, set theory, e.g. the system ZF of Zermelo and Fraenkel, is usually conceived as a reconstruction of pure mathematics with no physics coming into it. However, it is well known that Zermelo has considered a set theory involving *urelements*, as he called them, that could be identified with physical objects. In the following I shall use a modern version of Zermelo's theory as it can be found in Suppes' book of 1960.<sup>8</sup>

---

<sup>7</sup> The same answer would have to be given with respect to Sneed's theoretical/non-theoretical dichotomy, cf. J. D. Sneed, *The logical structure of mathematical physics*, Dordrecht, 1971.

<sup>8</sup> P. Suppes, *Axiomatic Set Theory*, Princeton, N. J., 1960.



Apart from the logical operations in the proper sense we would then have set theoretical membership as a logical relation, neutral with respect to the distinction between mathematical and physical entities. I am a member of the collection of people now present in this room in exactly the same sense in which 3 is a member of the set of prime numbers. The distinction between mathematical and physical objects can be made with the help of the axiom of foundation. From this axiom it follows that every descending chain

$$\dots \in x_2 \in x_1 \in x_0$$

is finite, the last object being either the empty set or an urelement.  $x_0$  is called *mathematical* if all its chains terminate in the empty set, *physical* if all the final elements are urelements and *mixed* in the rest of the cases. With the help of the predicates thus defined we can restrict quantifiers to mathematical or to physical objects and in this way distinguish between sentences being about mathematical objects, physical objects and about both. The axioms restricted to mathematical objects follow from the original ones, and we thus recover the usual version of set theory. But our generalization is now better prepared for the kind of investigation that we want to conduct.

Our next question is: how is a physical theory to be conceived within this set theoretical scheme? Very roughly, the primary semi-theory is assumed to be a species of structures  $\Sigma^1$  in the sense of Bourbaki.<sup>9</sup> The secondary semi-theory, *if it has axioms*, is also a species of structures  $\Sigma$ , *deduced* from the first, i.e.

$$(4.1) \quad \Sigma^1(a^1) \vdash_{\tau} \Sigma(\tau(a^1))$$

and, if possible, conservatively deduced, i.e. we have also the inverse implication

$$(4.2) \quad \Sigma(\tau(a^1)) \vdash_{\tau} \Sigma^1(a^1)$$

Here  $\tau$  are terms, intrinsic for  $\Sigma^1$ , and defining the fundamental embedding

$$(5) \quad \rho(\alpha(a)) \equiv \alpha(\tau(a^1))$$

of the secondary into the primary language. In general there are no secondary axioms but only a language created by object constants  $a$  and an implication  $\vdash$  which here is but the set theoretical implication  $\vdash_{\tau}$  of the primary semi-theory restricted to the secondary language. This is the situation as we encounter it in the examples from physics. Most textbooks for quantum mechanics, for instance, are proud to introduce complex Hilbert space – abstract or concrete – in

---

<sup>9</sup> N. Bourbaki, *Elements of Mathematics: Theory of Sets*, Paris, 1968. Ch. IV

the first place. But it is only its self-adjointed operators and its 1-dimensional subspaces that correspond to physical entities: to observables and states respectively. With the Hilbert space axioms we are offered typical primary axioms, and secondary axioms are only introduced as consequences of the former in the sense of (4.1).<sup>10</sup> Here, as in general, the main problem is the *existence* of the secondary axioms satisfying also (4.2).

To fill in some details I first draw your attention to the structures

$$(6) \quad a \equiv (X, s), \quad a^1 \equiv (X^1, s^1)$$

talked about in the two languages. Their elements may be of any of the three kinds introduced previously: physical, mathematical and mixed. However, the *s*-terms in (6) are assumed to be typified by the *X*-terms – the base terms, and this reduces their possibilities of their being of those kinds. For typification means that an *s*-set is an element of a set constructed from the base sets by iterating the operations of taking the power set of a cartesian product. The base sets *X* are assumed to be pure: physical or mathematical. Then a set typified by a mathematical base sets alone is again mathematical. The typification, for example, of the distance function *d* in physical geometry is

$$(7.1) \quad d \in \text{Pow}(M^2 \times \mathbb{R})$$

where *M* is physical space, and  $\mathbb{R}$  is the set of real numbers. By contrast the relation of betweenness *btw* has typification

$$(7.2) \quad \text{btw} \in \text{Pow}(M^3)$$

*d* is mixed whereas *btw* is physical. Therefore, if (7.1) and (7.2) appear in the primary and secondary axioms of a semi-theory decomposition of geometry then this would amount to an elimination of a mathematical object and, while observing the conservation requirements for  $\rho$ , be a step of reaching the goal of purely physical axiomatization.

The second and more difficult part of our problem concerns the *axioms proper*. Suppose we axiomatize euclidean distance geometry by requiring that there exists a coordinate system for space *M* that carries distance *d* into the well known arithmetically defined euclidean distance. This would be a most typical primary axiom: We say what we want to say about a physical structure by taking a loan from a much richer arithmetical structure: we require the former to be isomorphic to a fragment of the latter. Now contrast this axiom, for

---

<sup>10</sup> It is interesting to see how the introduction of Hilbert space is postponed in attempts to obtain physical axiomatizations. See G.W. Mackey, *Mathematical Foundations of Quantum Mechanics*, New York, 1963, Ch. 2–2.

instance, with the one about betweenness, saying that betweenness is a symmetrical relation with respect to the outer arguments. Obviously, this is precisely the kind of thing that we want a physical or, for that matter, a geometrical axiom to say. Besides logical operations and truly descriptive terms, besides logics and physics, no third subject matter enters the stage. If the whole language and, in particular, all axioms of geometry could be reformulated in this manner we would have obtained a secondary axiomatization at its best and could forget about the primary one.

It is well known that in the case of Euclidean geometry we *can* have an axiomatization of this kind.<sup>11</sup> Moreover, for nearly two thousand years geometry was treated according to this purely geometrical or synthetic method without any alternative in sight. It was only in the 17<sup>th</sup> century that the new analytical way of doing things began to evolve and indeed to dominate not only geometry but also the development of the new physics. Eventually the Euclidean tradition was crowned in our century by the work of Hilbert and Tarski leading to categorical geometrical axiom systems meeting all requirement of modern mathematical logic. Even within our present framework which still includes set theory, this achievement can be viewed as a classical solution of our main problem for ordinary geometry.

It is quite simple to see what demands we have to meet in order to make our secondary language, and with it, the secondary axioms physically acceptable in a general sense. We have already seen that the secondary typification should contain no mathematical terms. The same is, of course, to be required of the axioms proper. Here the chief trouble comes from the bound variables. In a species of structures they are allowed to run over the whole set universe, and it can easily be seen that, if this is allowed in the secondary language, our main problem always has a trivial solution. At the same time it is obvious why such a liberalism is unacceptable. If we want to make the secondary axioms physically plausible we shall have to restrict the bound variables not only to physical objects but actually to the physical structure that is the referent of the secondary object constants  $a$ . If, finally, the primitive formulas of the secondary language are confined to formulas

$$\langle x_1, \dots, x_n \rangle \in X$$

we certainly have reached an acceptable language – let us call it a *type reduced language* – on this level of generality. At the same time this restriction is so severe that our reduction problem becomes very hard in many cases.

---

<sup>11</sup> A. Tarski, "What is Elementary Geometry?". In: *The Axiomatic Method*. Ed. L. Henkin et al., Amsterdam, 1959, 16–29.

This can be seen most impressively in the work done by Hartry Field.<sup>12</sup> It concerns classical field theories. There are physical laws for scalar fields that are invariant against linear transformations of the field values. In such a case we may say that neither a unit nor a zero point is intrinsically fixed for these fields. An example is temperature with respect to the law of heat conduction. The true subject matter of such a theory is *not* a scalar field with well determined values. Rather, it is a certain equivalence class of such fields, and one has to characterize them by more elementary objects in the same way as in geometry we succeed in characterizing distance functions modulo a positive factor by the relations of congruence and betweenness. The scalar field is replaced by a sort of congruence relation concerning quadruples  $x, y, u, v$  of points in space telling us whether the absolute field difference between  $x$  and  $y$  equals that between  $u$  and  $v$ . Likewise a betweenness relation involving three points  $x, y$  and  $z$  tells us whether the field value in  $y$  is between that in  $x$  and  $z$ . Then physically reasonable axioms are strong enough in order to show every structure satisfying them to be deducible from a scalar field obeying the original law.

### 5. Approaching Type Logical Reconstructions

The most impressive attempt to eliminate a primary semi-theory in favour of secondary axioms formulated in purely physical terms was the attempt to fulfil the so called v. Neumann program, i.e. the program to achieve a physical understanding of the mathematics of complex Hilbert space as used in quantum mechanics. It seems that a solution is now in sight.<sup>13</sup> But for us it is time to have a look, if only a brief one, at *the two other cases* of semi-theory decomposition that are of interest with respect to the application of mathematics in physics. The characteristic feature of our first case was that both semi-theories had set theory as their logico-mathematical foundation. A further case of interest is the one where this only holds for the primary semi-theory whereas the secondary one is based directly on first or second order logic or any finite type logic. In the third case we say definitely good-bye to set theory and both logics involved are finite type logics.

Now first a few words about the *second case*. In it, the fundamental mapping  $p$  has to translate the language of a finite type logic into the language of set theory. A natural way to achieve this is well known. In the usual exten-

<sup>12</sup> H. Field, *Science Without Numbers*, Princeton, N. J., 1980; ch. 6–8.

<sup>13</sup> G. Ludwig, *An Axiomatic Basis for Quantum Mechanics*, vol. 1: *Derivation of Hilbert Space Structure*, Berlin, 1985.

sional interpretation of a finite type logic the predicates are interpreted by sets according to their type: The predicate types correspond exactly to the typification of sets by means of the power-set-of-cartesian-product operations. A syntactical version of the usual extensional interpretation of a finite type language will therefore immediately give us an injective mapping  $\rho$  of finite type sentences into set-theoretical sentences, where the latter are essentially type reduced species of structures as introduced at the end of the preceding section. Thus the first order sentence

$$\forall x, \forall y, Rxy \rightarrow Ryx$$

has the translation

$$R \in \text{Pow}(X^2) \wedge \forall x, \forall y, x, y \in X \rightarrow \langle xy \rangle \in R \rightarrow \langle yx \rangle \in R$$

where  $X$  is the universe of discourse. The definition of  $\rho$  has, of course, to be qualified by a remark concerning the distinction between physical and mathematical objects, terms and sentences. If we assume our finite type languages to be many-sorted, it is natural to require that each sort is either physical or mathematical and that the base sets corresponding to the different sorts are of the same kind. This conservation at the bottom will then hold all the way up to higher types that are involved on both sides. As opposed to the previous, purely set-theoretical case, there will thus be neither loss nor gain of mathematical or physical terms in either direction. What undergoes a change this time is the general logico-mathematical apparatus.

This can be seen immediately if we look at our earlier postulates (3), connecting the logics and axioms via  $\rho$ . There is no problem with the axioms this time. If their problem cannot be solved we have our miracle already in the first, purely set-theoretical case. If it can be solved, what formerly was our secondary semi-theory may now become the primary one. Since its language is type reduced it is natural to choose the axioms  $A$  of the secondary semi-theory such that  $A_1 = \rho(A)$ , and the second lines of (3) are satisfied. It is different with the logics: Although the first line of (3.1) is always satisfied, the same is not true for the inverse implication. This can be seen by the following consideration that is confined to finite sets of premises  $\Gamma$ . The premise in (3.2) (first line) is equivalent to the semantical implication of  $\alpha$  by  $\Gamma$ , according to the finite type logic in question. Now, if the latter is first order then by the weak Gödel completeness theorem we get the conclusion in (3.2). In any higher order base, however, Gödel's incompleteness theorem makes this inference dubious. In other words, unless we succeed in re-axiomatizing a physical theory to become first order we cannot get rid of set theory *if we start with it*. And first order axiomatization seems to be the exception, not the rule in contemporary physics.

So why are we not happy without set theory? This question leads to the *third case* where only finite type logics are involved. It is true that, as I said before, in admitting set theory we have the least trouble in reconstructing higher level theories of modern physics as quantum mechanics and general relativity. However, as a matter of principle, I think, we can avoid set theory in favour of finite type logics. The most interesting representation mappings  $p$  are then again definition eliminating mappings as they were in the set-theoretical case and, therefore, many situations typical for this case will recur under the new assumptions. They concern shifts of particular physical and mathematical concepts from one semi-theory to the other in both directions.

But there are also more general results concerning the logics involved. Some results show that the use of higher order logics does not lend itself to the typical miracle I am trying to describe. For instance, the second order closure of first order logic is a conservative extension even if all predicative comprehension axioms are added.<sup>14</sup> Higher order logics are not invoked in order to strengthen first order logic. Rather they seem necessary because either we cannot escape physical structures of higher order or we want to make higher order statements on first order structures. A typical case of the latter kind, that leads to interesting decompositions, are axiom systems in which the only second order axioms are of the form

$$(8.1) \quad \forall P, \alpha[P]$$

where  $P$  is a  $n$ -ary first order variable and  $\alpha$  does not contain any further second order quantifications. If we replace these axioms by axiom schemata

$$(8.2) \quad \alpha[\phi]$$

where  $\phi$  is an arbitrary  $n$ -ary first order formula the resulting first order axiom system must not but can be conservatively represented in the original one. In particular, this will happen if the latter is complete. Axioms of geometrical continuity are a case in point.<sup>15</sup>

## 6. *Overladenness vs. Preestablished Harmony*

In this paper I began by quoting some statements of leading scientists telling us that the effectiveness of mathematics in the natural sciences is unreasonable and even miraculous. I argued that their attempts to explain this effectiveness had not been very successful because they had not sufficiently analysed the

---

<sup>14</sup> G. Takeuti, *Proof Theory*, Amsterdam, 1975. Ch. 3. Cor. 16.3 and 16.6

<sup>15</sup> See above n. 11.

phenomenon. I myself did not make an attempt to give an explanation. Rather I tried to analyse the situation with the aim of showing that we *really* have reasons to marvel at that effectiveness. In conclusion let me make one last effort to make my essential idea clear by giving you a mirror image of my argument. To this end I shall have to quote some remarkable passages from P.W. Bridgman's *The Nature of Physical Theory*.<sup>16</sup> With the new quantum mechanics in mind Bridgman describes the relation of the mathematics of quantum mechanics to the physically significant part of that theory in the following way:

In our elementary and classical theories we have become used to discarding perhaps one-half of the results of mathematics, [...], but here we retain only an infinitesimal part of the mathematical results, and except for a few isolated singular points relegate the entire mathematical structure to a ghostly domain with no physical relevance. A vivid appreciation of this situation will make it rather difficult, I believe, to maintain a conviction of the organic similarity of mathematics and physical experience, [...].

The view that Bridgman alludes to in the last sentence has been described by him earlier in the book in the following words:

The feeling that all the steps in a mathematical theory must have a counterpart in the physical system is the outgrowth, I think, of a certain mystical feeling about the mathematical construction of the physical world. Some sort of an idea like this has been flitting about [...] at least since the days of Pythagoras, and every now and then, [...], it bursts forth again like a crop of mushrooms after a rain, as in the recent fervid exclamation of Jeans that "God is a mathematician".

With this 'exclamation' we would be back with Kepler and Galileo if there were not Bridgman's own view on the matter. Here it is:

There would seem to be no necessity [...] that all mathematical operations should correspond to recognizable processes in the physical system. Nor is there any more any reason why all the *symbols* appearing in the fundamental mathematical equations should have their physical counterpart, [...].

All that is required of the theory is that it should provide the tools for calculating the behaviour of the physical system, and it is capable of doing this if there is correspondence between those aspects of the physical system which it engages to reproduce and *some* of the results of the mathematical manipulations.

---

<sup>16</sup> P. W. Bridgman, *The Nature of Physical Theory*, Princeton, N. J., 1936 The quotations are from pp. 116 f, 67, 66, 65 in this order.

You will have realized that on a not too fine scale Bridgman gave the same analysis of the relation in question that I have given. The consequence, however, that he draws from this is markedly different, if not opposite. He denies any reason to be disappointed by the failure of the 'organic similarity' view simply because mathematics was made by man rather than by God. Now this may very well be the case. But it is compatible with the consequence I have drawn: The overladdenness of physics with mathematics is strong evidence that physical theories are non-conservatively embedded in mathematics. And this situation appears to me to be *more* sophisticated than the view of a preestablished harmony. Moreover, why should man not be able to conceive of such a sophistication. Let me bring this lecture to an end by quoting Einstein who said: "...all our science, measured against reality, is primitive and childlike [...] and yet it is the most precious thing we have".<sup>17</sup>

---

<sup>17</sup> B. Hoffmann, *Albert Einstein. Creator and Rebel*, The Viking Press, 1972, p. VII.



# The Status of Set-theoretic Axioms in Empirical Theories

HEINZ-JÜRGEN SCHMIDT (Osnabrück)

## 1. Introduction

1.1. In an essay concerning the foundations of mathematics Lakatos (1978) has argued that (meta-)mathematics should be conceived as a "quasi-empirical" theory. He even raises the following question:

"Are we going to arrive, tracing back problemshifts through informal mathematical theories to empirical theories, so that mathematics will turn out in the end to be *indirectly empirical*, thus justifying Weyl's, von Neumann's and – in a certain sense – Mostowski's and Kalmar's position?"

In my paper I will try to explore a particular aspect of this question: in which sense and to which extent could a part of (meta-)mathematics, namely the set-theoretical axioms, obtain an empirical meaning? Unfortunately, I cannot present a definite answer, but I will arrive at a distinction between the empirical and the non-empirical part of set theory which probably proves right if some strong restrictions concerning the language of empirical theories are imposed.

1.2. The axioms postulated in an axiomatization of a physical<sup>1</sup> theory usually can have a different methodological status: they could be empirical laws in a narrow sense, or implicit definitions of (physical or auxiliary) concepts, or rules delimitating the range of intended application, or idealizations of different kinds. Moreover, an axiom could be of mixed character or construed in different ways. One aim of the reconstruction of physical theories is to analyze and clarify the status of the axioms.

It is not quite obvious how to extend such an analysis to set-theoretic axioms. A first problem is that there are at least two different approaches concerning the way set theory comes into contact with an empirical theory.

---

<sup>1</sup> Though I am mainly interested in *physical* theories, I cannot see any reason to restrict my investigation in this way. Hence the title is formulated more generally.

- (i) The first approach is based on the theory concept of mathematical logic, according to which a theory consists of a *language*, non-logical *axioms*, and a *calculus of formal proofs*, given for example by *logical axioms* and *inference rules*, cf. Shoenfield (1967), or by *sequential rules*, cf. Ebbinghaus, Flum and Thomas (1986). *Models* of the theory are then defined in terms of (usually informal) set theory.
- (ii) The second theory concept considers an extension of set theory by individual constants and additional axioms, defining a *set-theoretic predicate* or, if additional conditions are satisfied, a *species of structures*  $\Sigma$ , cf. Bourbaki (1968).

There are certain connections between the two theory concepts.

First, almost every theory of kind (i) gives rise to a species of structure. This will be illustrated by a simple example. Let  $T$  be a first order one-sorted theory with a single binary relational symbol  $R$  and the axioms of an order relation. Then the corresponding species of structure will be specified by constants  $A, r$  where  $A$  is the single base set, the typification axiom reads " $r \in \mathcal{P}(A \times A)$ " and the proper axiom will be the translation of the axioms of order.

Conversely, set theory itself can be reconstructed as a first order theory of kind (i) with a single kind of objects called "sets" and a relational symbol " $\in$ ", which together with certain axioms constitutes the so-called Zermelo-Fraenkel (ZF)-system, cf. for example Drake (1974). The extension of ZF to a species of structure  $ZF\Sigma$  will not injure this characterization and thus one may again consider models of  $ZF\Sigma$  in the previous sense.

Both theory concepts have been used for reconstructing empirical theories and, beyond this, as requisites for a meta-theory of science. Set-theoretical predicates are used by Suppes (1970) and Sneed (1971). The idea to use species of structures for reconstructing physical theories and the most important example concerning quantum mechanics is due to Ludwig (1978, 1985, 1987, 1990) and further investigated by Scheibe (1978, 1982, 1986). It has also recently been incorporated into the structuralistic approach in Balzer, Moulines, and Sneed (1987). I will mainly focus on Ludwig's approach, but my remarks *grano cum salis* will also apply to the structuralistic approach despite some differences in the use of species of structures between the two doctrines.

1.3. If a physical theory has been cast into the form  $ZF\Sigma$ , the proper axioms of  $\Sigma$  appear on a par with the ZF-axioms, the only difference being the "universal" character of the latter. "Universal" means that the ZF-axioms quantify over the universe of all sets and invariantly occur in any physical theory, whereas the  $\Sigma$ -axioms typically quantify over  $\Sigma$ -specific sets and change from theory to theory. The second aspect of universality does not

necessarily hinder an empirical interpretation of the  $ZF$ -axioms. It could be historically accidental in the sense that serious alternatives to some, or all,  $ZF$ -axioms may today be equally poorly investigated<sup>2</sup> just as the alternatives to Euclidean geometry 150 years ago.

The difficulties in analyzing the  $ZF$ -axioms are analogous to those classifying the status of the  $\Sigma$ -axioms alone. Only a part of the terms of the theory can be physically interpreted, the remaining terms are "purely mathematical" and typically needed to define a part of the interpreted terms. But usually both kinds of terms occur in the axioms of the theory and give them the mixed character mentioned in section 1.2. Ludwig in his (1978) developed a technique to reformulate physical theories in such a way that only interpreted terms occur, thereby "purifying" a theory from superfluous mathematics. His program is even more ambitious: in the reformulated theory, called *axiomatic basis*, all theoretical terms should be eliminated. But this point will be controversial and would need further explanation. So I will only assume that the proper terms of the theory, i.e. the individual constants of  $ZF\Sigma$ , can be considered as empirical and the proper axiom of  $\Sigma$  as an empirical law (in a wide sense).

1.4. Nevertheless, in  $ZF\Sigma$  we have a mixed vocabulary of empirical and set-theoretical terms, and since the empirical terms are also sets, the  $ZF$ -axioms refer to both kinds of terms. In principle, it is clear what the "empirically effective" part of  $ZF$  should be; let us call it  $\tilde{ZF}$ : it consists of all logical consequences of  $ZF$  that are expressible in the empirical vocabulary. But there is a real problem to write down  $\tilde{ZF}$  as a finite number of axioms, not *à la* Craig's theorem, and even to decide in which language  $\tilde{ZF}$  should be formulated.

Fortunately, there exists a paper of Scheibe (1986), where he indeed not solved these elimination problems, but gave them a more precise meta-mathematical formulation in the context of a discussion of Field's book (1980).

Scheibe considers a many-sorted finite order theory  $T$  "underlying"  $ZF\Sigma$  in the sense, that the different sorts of variables of  $T$  correspond to different base sets of  $ZF\Sigma$  and the relation and function symbols of  $T$  correspond to the typified sets  $S_1, \dots, S_k$  of  $ZF\Sigma$ . If the proper axiom of  $\Sigma$  is, in a certain sense, *typified*, it can be re-translated into an axiom of  $T$ . The conjectured *main theorem* of Scheibe (1986) then states that  $ZF\Sigma$  is the analogue of a conservative extension of  $T$ .

---

<sup>2</sup> To be sure,  $ZF$  set theory is not categorical and there exist different possibilities to add axioms to  $ZF$ , for example concerning large cardinals. Jech (1973) mentions the *axiom of determinateness* that implies the countable axiom of choice, but contradicts the general axiom of choice.

We will reformulate Scheibe's conjecture in a bit more detailed manner and prove it for the special case of first order theories. Admittedly, this case is hardly interesting for real life empirical theories, but it will serve as our starting point for tackling the case of higher order theories. A partial solution of the problem addressed in the title will be sketched in the last paragraph.

## 2. Scheibe's Conjecture for First Order Theories

2.1. It appears as a natural choice to consider for  $T$  a many-sorted language, because empirical theories will distinguish between different sorts of empirical objects, as e.g. particles, events, fields. From a meta-mathematical point of view this choice is not a severe restriction because many-sorted languages can be reduced to one-sorted ones by using a finite number of predicates  $T_n$  with the interpretation "... is of sort  $n$ ". It is, however, preferable to distinguish the  $T_n$  from the other predicates in order to obtain a unique species of structure  $\Sigma(T)$ , cf. 2.3.

2.2. So we will consider  $N$  sorts of objects,  $I$  relations and  $K$  individual objects.<sup>3</sup> The language  $\mathcal{L}(T)$  of the theory  $T$  will be characterized by the signature  $\tau = (\Phi, \kappa)$ , where  $\Phi$  and  $\kappa$  are mappings

$$\Phi : I \times N \rightarrow \mathbb{N}, \quad \kappa : K \rightarrow N$$

(compare Potthoff (1981)). The alphabet of  $\mathcal{L}(T)$  will consist of

- (i) for each sort <sup>4</sup>  $n \in N$  a countable set of variables  $V^{(n)}$  such that  $V^{(n)} \cap V^{(m)} = \emptyset$  for  $n \neq m$ ,
- (ii) the logical symbols  $\neg, \vee$ , and  $\exists$ , and brackets  $(, )$
- (iii) for each sort  $n \in N$  a relational symbol<sup>5</sup>  $E_n$  of equality,
- (iv) for each  $i \in I$  a relational symbol  $R_i$ ,
- (v) for each  $k \in K$  an individual constant  $c_k$ .

Variables of different sorts will be distinguished by subscripts, e.g.  $v_n$  for a variable of sort  $n$ . The terms of  $\mathcal{L}(T)$  are either variables or individual constants. The individual constant  $c_k$  will be said to be of sort  $\kappa(k)$ , so sort  $(t)$

<sup>3</sup> For sake of simplicity, I will omit the functional symbols. They are in principle dispensable, but a full-fledged theory should admit the possibility to add new defined functional symbols.

<sup>4</sup> Here I consider natural numbers as finite ordinals, hence each natural number is identified with the set of all smaller natural numbers, starting with  $0 = \emptyset$ .

<sup>5</sup> The introduction of "typified equality" symbols is crucial for theorem T3.

will be defined for all terms  $t$ . *Atomic formulas* of  $\mathcal{L}(T)$  will be defined as follows:

- (i) If  $s, t$  are terms of sort  $n$ , then  $E_n st$  will be an atomic formula,
- (ii) If  $t_1 \dots t_L$  are terms and  $i \in I$  such that  $\ell \rightarrow \text{sort}(t_\ell)$ ,  $\ell \in L$ , is monotonely increasing and the number of terms in  $t_1 \dots t_L$  of sort  $n$  is equal to  $\Phi(i, n)$ , then  $R_i t_1 \dots t_L$  will be an atomic formula.

These conditions insure that relational symbols are only fed with terms of suitable sorts. *Compound formulas* will be defined as usual. Of course, *substitution* is only possible for terms of the same sort. The *sequential rules* governing formal proofs can be adopted from Ebbinghaus, Flum, and Thomas (1986) with the only modification that the two rules of identity are to hold for each sort  $n \in N$ . Finally a finite set  $\mathcal{A}$  of closed formulas (*axioms*) is singled out in  $T$ . Theorems are formulas  $F$  which can be proved under the premises  $\mathcal{A}$ , in symbols:  $\mathcal{A} \vdash F$  or, synonymously,  $T \vdash F$ .

The *semantics* of  $T$  can be defined as usual, see e.g. Kreisel and Krivine (1971) for the semantics of many-sorted languages. Thus a *structure* of  $\mathcal{L}(T)$  is a  $N+I+K$ -tuple of sets  $\mathcal{U} = \langle \mathcal{U}_1, \dots, \mathcal{U}_N, \mathcal{Q}_1, \dots, \mathcal{Q}_{I+K} \rangle$ , where  $\mathcal{Q}_i \subset \prod_n \mathcal{U}_n^{\Phi(i,n)}$  and  $\mathcal{Q}_{i+k} \in \mathcal{U}_{\kappa(k)}$ . Further, the *validity* of a formula  $F$  in a structure  $\mathcal{U}$ ,  $\mathcal{U} \models F$ , is defined as usual.<sup>6</sup> For example, " $E_n st$ " is valid in a structure iff the sets corresponding to the terms  $s$  and  $t$  are equal. Finally, a *model* of  $T$  is a structure  $\mathcal{U}$  in which the axioms  $\mathcal{A}$  of  $T$  are valid.

2.3. To each theory  $T$ , as described above, we will assign a *species of structure*  $\Sigma(T)$ , which will be, so to speak, the theory  $T$  in a set-theoretic guise. Following Scheibe (1986), we will define the species of structure as an extension  $ZF\Sigma(T)$  of Zermelo-Fraenkel set theory  $ZF$  in such a way that  $T$  and  $ZF\Sigma(T)$  have essentially the same models. To this end we extend  $ZF$  by the usual definitions of logical and set-theoretic abbreviation symbols and additional individual constants  $A_1, \dots, A_N$  and  $S_1, \dots, S_{I+K}$  and axioms  $\mathcal{T}\mathcal{A}^\sim$ . Intuitively, the  $A_n$  are abstract sets of individuals of different sorts  $n \in N$ , and  $S_1, \dots, S_{I+K}$  are set-theoretic equivalents for the relations and individuals denoted by  $R_i$ ,  $i \in I$ , and  $c_k$ ,  $k \in K$ . This motivates the formulation of the *typification axioms*  $\mathcal{T}$ :

$$S_i \in \mathcal{P}(\prod_n A_n^{\Phi(i,n)}), i \in I, \text{ and } S_k \in A_{\kappa(k)} \quad k \in K.$$

In order to specify  $\mathcal{A}^\sim$  we have to define a translation procedure  $\sim$  of  $\mathcal{L}(T)$ –

<sup>6</sup> The validity of formulas with free variables of course depends on the values of these variables in the given structure. Since we are mainly interested in closed formulas, we need not consider this additional dependence.

formulas into  $\mathcal{L}(ZF\Sigma(T))$ -formulas. The point is that the information about the sort of a term which is inherent in a  $\mathcal{L}(T)$ -formula  $F$  has to be conserved in  $F^\sim$ . Let  $V$  denote the countable set of variables in  $\mathcal{L}(ZF)$  and consider an injective mapping, again denoted by  $\sim$  and written as a superscript

$$\sim : \bigcup_n V^{(n)} \rightarrow V.$$

The mapping  $\sim$  will be extended to all terms of  $\mathcal{L}(T)$  by setting

$$c_k^\sim \equiv S_{I+k}, k \in K.$$

For atomic formulas and its negations, we will define ( $[\neg]$  is optional):

- (i)  $([\neg] E_n st)^\sim \equiv (s^\sim \in A_n \wedge t^\sim \in A_n \wedge [\neg] s^\sim = t^\sim)$
- (ii)  $([\neg] R_i t_1 \dots t_\ell)^\sim \equiv (\langle t_1 \dots t_\ell \rangle \in (\prod_n A_n^{\Phi(i,n)}) \wedge [\neg] \langle t_1 \dots t_\ell \rangle \in S_i).$

For the remaining compound formulas we will define recursively:

- (iii)  $(F \vee G)^\sim \equiv F^\sim \vee G^\sim$  and  $(\neg (F \vee G))^\sim \equiv (\neg F)^\sim \wedge (\neg G)^\sim$ ,
- (iv)  $(\exists x_n F(x_n))^\sim \equiv (\exists x (x \in A_n \wedge F^\sim(x)))$ ,
- (v)  $(\neg \exists x_n F(x))^\sim \equiv (\forall x (x \in A_n (\neg F)^\sim(x)))$ ,
- (vi)  $(\neg \neg F)^\sim \equiv F^\sim$ .

If  $S$  is a set of formulas of  $\mathcal{L}(T)$ , we will set

$$S^\sim \equiv \{F^\sim \mid F \in S\}.$$

Then the set  $\mathcal{A}^\sim$  of the translated axioms of  $T$  will be the set of *proper axioms* of  $ZF\Sigma(T)$ , which completes the definition of the species of structure  $ZF\Sigma(T)$ .

2.4. Let  $\mathcal{U} \equiv \langle |\mathcal{U}|, \hat{e} \rangle$  be a fixed model of  $ZF$ . An expansion

$$\mathfrak{M} \equiv \langle |\mathcal{U}|, \hat{e}; \mathfrak{U}_1, \dots, \mathfrak{U}_N, \mathfrak{Q}_1, \dots, \mathfrak{Q}_{I+K} \rangle$$

of  $\mathcal{U}$  which is a model of the typification axioms  $\mathcal{T}$  of  $ZF\Sigma(T)$  will be called a *potential  $\mathcal{U}$ -model* [of  $ZF\Sigma(T)$ ], and  $\mathfrak{M}$  will be called a  *$\mathcal{U}$ -model* [of  $ZF\Sigma(T)$ ] if additionally the proper axioms  $\mathcal{A}^\sim$  are valid in  $\mathfrak{M}$ . This definition obviously parallels the structuralistic terminology. Analogously, a structure

$$\mathfrak{U} \equiv \langle \mathfrak{U}_1, \dots, \mathfrak{U}_N, \mathfrak{Q}_1, \dots, \mathfrak{Q}_{I+K} \rangle$$

of  $\mathcal{L}(T)$  with sets taken from  $\mathcal{U}$  will be called a  *$\mathcal{U}$ -model* of  $T$ . If  $\mathfrak{U}$  results from  $\mathfrak{M}$  by omitting  $|\mathcal{U}|, \hat{e}$  as above, we will write

$$\mathfrak{U} = \mathfrak{M} \sim.$$

The following theorem follows in a straight forward manner from the definition of  $\models$  and  $\sim$ :

T1: Let  $\mathfrak{M}$  be a potential  $\mathcal{U}$ -model of  $ZF\Sigma(T)$  and  $F$  a closed formula of  $\mathcal{L}(T)$ . Then  $\mathfrak{M} \models F \sim$  iff  $\mathfrak{M} \sim \models F$ .

T1 implies:

T2: Let  $\mathfrak{M}$  be a potential  $\mathcal{U}$ -model of  $ZF\Sigma(T)$ . Then  $\mathfrak{M}$  is a  $\mathcal{U}$ -model of  $ZF\Sigma(T)$  iff  $\mathfrak{M} \sim$  is a  $\mathcal{U}$ -model of  $T$ .

The reader who is familiar with the Bourbaki concept of a species of structure will wonder whether the above defined proper axioms  $\mathcal{A} \sim$  will be *transportable* as required by Bourbaki. This condition can be reformulated as follows:

A closed formula  $\phi$  of  $\mathcal{L}(ZF\Sigma(T))$  will be called *transportable* iff for every two isomorphic potential  $\mathcal{U}$ -models  $\mathfrak{M}^{(1)}$  and  $\mathfrak{M}^{(2)}$  of  $ZF\Sigma(T)$  we have:

$$\mathfrak{M}^{(1)} \models \phi \text{ iff } \mathfrak{M}^{(2)} \models \phi.$$

It should be stressed that the notion of an *isomorphism* between  $\mathcal{U}$ -models has to be defined analogously to Bourbaki(1968), i.e. only with respect to the  $\mathfrak{M} \sim$  – part of  $\mathfrak{M}$ , not with respect to  $\hat{\epsilon}$ . Otherwise every closed formula would be transportable by virtue of the isomorphism lemma (cf. Ebbinghaus, Flum, and Thomas (1986), 5.5, for one-sorted languages). It turns out that the condition of transportability, which is also of eminent importance in physical theories, cf. Scheibe (1982), is automatically satisfied for  $\mathcal{A} \sim$ :

T3: For every closed formula  $F$  of  $\mathcal{L}(T)$ ,  $F \sim$  will be transportable.

Proof: Let  $\mathfrak{M}^{(1)}$  and  $\mathfrak{M}^{(2)}$  be isomorphic potential models, then  $\mathfrak{M}^{(1)} \sim$  and  $\mathfrak{M}^{(2)} \sim$  will be isomorphic  $\mathcal{L}(T)$ -structures. By the isomorphism lemma,  $\mathfrak{M}^{(1)} \models F \sim$  iff  $\mathfrak{M}^{(2)} \models F \sim$ , hence, applying T1,  $\mathfrak{M}^{(1)} \models F \sim$  iff  $\mathfrak{M}^{(2)} \models F \sim$ . Now we can formulate and prove Scheibe's conjecture:

T4: Let  $F$  be a closed formula of  $\mathcal{L}(T)$ . Then  $T \vdash F$  iff  $ZF\Sigma(T) \vdash F \sim$ .

Proof: (i) For proving the if-part let  $\mathfrak{M}'$  be a  $\mathcal{U}$ -model of  $T$ . It can be written as  $\mathfrak{M}' = \mathfrak{M} \sim$ , where  $\mathfrak{M}$  is a  $\mathcal{U}$ -model of  $ZF\Sigma(T)$ . Because the sequential calculus is correct, we have  $\mathfrak{M} \models F \sim$ . By T1 we conclude  $\mathfrak{M}' \models F$  and, since  $\mathfrak{M}'$  was an arbitrary  $\mathcal{U}$ -model and  $T$  is complete,  $T \vdash F$ .

(ii) For the only-if-part a possible strategy would be to show that any proof of  $F$  in  $T$  can be translated into a proof of  $F \sim$  in  $ZF\Sigma(T)$ . This seems plausible and I do not want to go into the details concerning sequential rules. Let me only mention the following difficulty, which makes the proof not completely trivial: Let

$$\rho \equiv \Gamma_1 F_1 / \dots / \Gamma_m F_m // \Gamma F$$

be a sequential rule<sup>7</sup> for  $\mathcal{L}(T)$ -formulas (written "horizontally"). Then the "pointwise" translation

$$\rho \sim \equiv \Gamma_1 \sim F_1 \sim / \dots / \Gamma_m \sim F_m \sim // \Gamma \sim F \sim$$

in general will not be a correct sequential rule for  $\mathcal{L}(ZF\Sigma(T))$ -formulas. One counter-example would be the rule of contradiction,

$$\Gamma \neg F G / \Gamma \neg F \neg G // \Gamma F,$$

where  $F$  and  $G$  contain free variables. But if  $\rho \sim$  is further transformed into a rule  $\rho^*$  by removing all typifications of free variables from the conclusions  $F_m \sim$  and putting them into the premises  $\Gamma_m \sim$ , it can be shown that  $\rho^*$  will be correct and the proof of a closed formula in  $T$  can be transferred to  $ZF\Sigma(T)$ .

### 3. Are Higher Order Theories Necessary?

3.1. The result T4 of chapter 2 does not bear much upon the status of set-theoretic axioms in empirical theories, because there are no first order axiomatizations of interesting empirical theories (except those already formulated in the framework of  $ZF$  theory). It will nevertheless turn out useful to closer examining the different reasons for this.

3.2. A first difficulty in confining oneself to first order theories arises, because often a sort of empirical objects is introduced at a higher set-theoretical level with respect to other sorts. For example, in geometry *lines* may be considered as certain sets of *points*. But this difficulty can be circumvented by using additional and independent sort relations as substitutes for set-theoretical relations. In the given example it would suffice to considering an *incidence relation* between points and lines (or higher-dimensional planes) and, if necessary, to formulating additional axioms which insure properties of the incidence relation analogous to those of " $\in$ ".

It is always possible to consider sorts of higher order level objects ("collections") and "incidence" relations in order to mimic a part of set theory within a first order theory, see Kreisel and Krivine (1971). The need for using higher order theories only arises if the intended interpretation of a sentence in a

<sup>7</sup> We will shortly explain the concept of sequential rules. A sequential rule is a finite sequence of *proof lines* with the following meaning: if the first  $n-1$  proof lines are legitimate, then also the last  $n$ -th proof line will be legitimate. A proof line is a finite sequence of formulas where the first  $k-1$  formulas are the premises and the last  $k$ -th formula is the conclusion.



model depends on quantifying over e. g. *all* subcollections of a given sort of objects. Thus, at the core, the difference between first and higher order theories is not a syntactical one but lies in the semantics of theories.

3.3. The need for quantifying over sets of relations may also occur due to the treatment of theoretical terms by so-called Ramsey sentences. An axiom of the kind "there exists a Lagrangian  $L : \mathcal{K} \rightarrow \mathbb{R}$  such that the equations of motion assume the form ..." could not be formulated at the first order level. We will, however, assume that such Ramsey sentences can be eliminated in favour of representation theorems. There are various examples for such an elimination.<sup>8</sup>

By analyzing these examples one obtains a more profound reason for using higher order theories: Axiomatizations usually aim at characterizing their models or at least some components of their models *categorically*, i.e. uniquely up to isomorphisms. As examples we only mention the separable complex Hilbert space in quantum mechanics or spacetime structures affinely isomorphic to  $\mathbb{R}^4$ . It is well known<sup>9</sup> that first order theories are unable to categorically characterize infinite structures as  $\mathbb{N}$ , let alone  $\mathbb{R}^4$ .

On the other hand, it has been pointed out by Ludwig (1978) that there is no empirically relevant difference between a set  $X$  (of descriptions of possible empirical objects or relations) and its *completion*  $\hat{X}$  with respect to a suitable uniform structure. The completion of an empirical set adds to the set a lot of ideal elements without postulating the possibility to constructing or finding new empirical objects, because the correspondence between mathematical and physical objects is at most an approximate one. Therefore the categoricity argument for higher order theories is not as stringent as it looks at first sight. Although the empirical laws are usually formulated at the "completion level", it should in principle be possible to find an empirically equivalent, albeit more complicated reformulation at the "non-completion level", cf. Schmidt (1981). If this would be true the most important role of set theory would be to accomplish a form of the empirical theory which is easier to handle without enlarging its empirical content.

3.4. We are then again faced with the question: what is the status of set theory in empirical theories at the non-completion level? As mentioned above, this form of axiomatization can be very complicated and no explicitly worked out example is known to me. However, the axiomatization of "rigid-body-

---

<sup>8</sup> In the terminology of Ludwig (1990) these are examples, where the *auxiliary theoretical terms* have been eliminated. Particularly, a simple axiomatic basis of a theory will be of this form.

<sup>9</sup> Cf. Tarski's cardinality theorem, Shoenfield (1967) 5.3.

geometry" given in my (1979) comes close to it. It could be reformulated as a theory  $T$  with two sorts of objects, *regions* and *transports*, and two relations, " $a \sqsubset b$ " for the inclusion of regions and " $\text{Op}(\tau, a, b)$ " for the case in which the transport  $\tau$  operates on the region  $a$  and yields a new (congruent) region  $b$ . A categorical representation theorem for the completed models of the theory is derived by imposing axioms R1 to R8, among which R1 to R5 and R7 are in the non-completion form. I will take it for granted that R6 and R8 could also be cast into this form and use this example of a geometric theory in order to further investigate the necessity to use higher order theories.

It turns out that the theory  $T$  even after the reformulation according to 3.2 cannot be retranslated into a first order theory because some of its axioms quantify over *finite* subsets of regions or transports. For example, R4 postulates some kind of Archimedean property, namely that every region can be covered by a finite number of arbitrarily small, congruent regions. As mentioned in 3.2 we could consider in  $T$  new sorts of "collections" of regions, resp. transports, and corresponding "incidence" or "membership relations". But it is not possible to adequately retranslate R4 into a first order theory, because it is impossible to first order characterize the family of finite collections in such a way that it corresponds exactly to the family of finite subsets in every  $\mathcal{U}$ -model of  $T$ .<sup>10</sup>

Such a characterization is well possible in higher order theories, but these lack completeness which was crucial for our proof of T4. Thus, in higher order theories we cannot a priori exclude empirical consequences  $F \sim$  of set-theoretic axioms in  $ZF\Sigma(T)$  which cannot be derived in  $T$  alone. It is not likely that this defect can be remedied by imposing new axioms in  $T$ , and it seems inevitable to abandon the framework of first order at least minimally.

#### 4. Tools From Infinitary Logics

4.1. As a way out of the dilemma sketched in the foregoing section, I propose to extend  $\mathcal{L}(T)$  by certain formulas occurring in infinitary languages  $\mathcal{L}_{\omega_1, \omega}$ . These are languages allowing for infinite disjunctions  $\bigvee S$ , where  $S$  is a countable set of formulas. If the sequential calculus of  $\mathcal{L}_{\omega_1, \omega}$  is equipped with suitable rules involving  $\bigvee$ ,  $\mathcal{L}_{\omega_1, \omega}$  is still complete,<sup>11</sup> at the price of admitting infinitely long proofs. Obviously, the finiteness of a collection  $X$  can be

<sup>10</sup> Assume that *fin* is a predicate expressing finiteness. Then the set of formulas *fin* ( $X$ ),  $|X| > 1$ , ... ,  $|X| > n$ , ... has no model, but any finite subset has a model. This contradicts the compactness theorem, cf. Shoenfield (1967) 5.1.

<sup>11</sup> Cf. Potthoff (1981), Satz 7.5.

expressed in this language by

$$\text{fin}(X) \equiv \bigvee \{F_n(X) \mid n \in \mathbb{N}\}$$

where

$$F_n(X) \equiv \exists x_1, \dots, \exists x_n \forall x (x_1 \neq x_2 \wedge x_1 \neq x_3 \wedge \dots x_n \neq x_{n+1} \wedge x \in X \leftrightarrow (x=x_1 \vee \dots \vee x=x_n)).$$

For easier readability we wrote " $\neq$ " for the negation of the equality relation and " $\in$ " for the membership relation in  $T$  (not  $ZF$ !) and used only one sort of variables. Finiteness is reflected by  $\mathcal{U}$ -models  $\mathfrak{M}_\sim$  of  $T$ , because  $\mathfrak{M}_\sim \models \bigvee \{F_n(X) \mid n \in \mathbb{N}\}$  iff, by definition, there is some  $n \in \mathbb{N}$  such that  $\mathfrak{M}_\sim \models F_n(X)$ . In order to insure the converse, namely that in every  $\mathcal{U}$ -model  $\mathfrak{M}_\sim$  of  $T$  every finite subset corresponds to a finite collection, we would have to impose a "finite subset axiom scheme" in  $T$ . Additionally, we have to consider translations of " $\text{fin}(X)$ " into a suitable  $\mathcal{L}(ZF\Sigma(T))$ -formula " $\text{fin}^\sim(X^\sim)$ ", e.g. a formula saying that each injection  $f: X^\sim \rightarrow X^\sim$  will be surjective. Then, in addition to T1, we have the result:

$$\mathfrak{M}_\sim \models \text{fin}(X) \text{ iff } \mathfrak{M}_\sim \models \text{fin}^\sim(X^\sim).$$

Moreover, T1 could be extended to closed formulas quantifying over finite collections. Further, it will be necessary to consider additional formulas from  $\mathcal{L}_{\omega_1\omega}$ . In the above example of rigid-body-geometry we had to employ the  $\sqsubset$ -supremum over an arbitrary finite collection of regions. A formula containing such a supremum term could also be expressed in  $\mathcal{L}_{\omega_1\omega}$  and translated into a  $\mathcal{L}(ZF\Sigma(T))$ -formula containing a recursively defined supremum term. Again it seems plausible that T1 extends to such formulas and hence the part of T4, i.e. of Scheibe's conjecture, will be provable for those formulas.

We will not work out this in detail, because at the moment we cannot give a concise characterization of a sub-language  $\mathcal{L}_{\text{emp}}$  of  $\mathcal{L}_{\omega_1\omega}$  which is rich enough to express empirical axioms, allows adequate translations into  $\mathcal{L}(ZF\Sigma(T))$  and will satisfy Scheibe's conjecture. The above example, however, indicates that the extension of  $\mathcal{L}(T)$  by formulas of the form  $\bigvee S$ , where  $S$  is recursively enumerable would be a good candidate for  $\mathcal{L}_{\text{emp}}$ .

4.2. If such an empirical language  $\mathcal{L}_{\text{emp}}$  would exist, it would only give an empirical meaning to that part of set theory dealing with finite sets or, as far as recursive definitions are employed, countable infinite sets. The other parts of  $ZF$  would have to be regarded as useful, but in principle superfluous elements of empirical theories.

### Acknowledgement

I am indebted to S. Feferman, J. Mosterin, F. Naishtat, and E. Scheibe for discussions on some topics of this paper, which were very helpful for me to develop a less naive view on logics.

### References

- W. Balzer / C. U. Moulines, / J. D. Sneed (1987), *An Architectonic for Science*, Reidel, Dordrecht.
- N. Bourbaki (1968), *Theory of Sets*, Herman, Paris.
- F. R. Drake (1974), *Set Theory*, North Holland, Amsterdam.
- H. D. Ebbinghaus / J. Flum / W. Thomas (1986), *Einführung in die mathematische Logik*, Wiss. Buchges, Darmstadt.
- H. Field (1980), *Science Without Numbers*, Princeton.
- T. J. Jech (1973), *The Axiom of Choice*, North Holland, Amsterdam.
- G. Kreisel / J. L. Krivine (1971), *Elements of Mathematical Logic*, North Holland, Amsterdam.
- I. Lakatos (1978), "A renaissance of empiricism in the recent philosophy of mathematics". In: J. Worrall / G. Currie (eds.), *Mathematics, Science and Epistemology, Philosophical Papers*, Vol.2 Cambridge UP, Cambridge.
- G. Ludwig (1978,1990), *Die Grundstrukturen einer physikalischen Theorie*, Springer, Berlin, 1st edn. 1978, 2nd edn. 1990.
- G. Ludwig (1985,1987), *An Axiomatic Basis for Quantum Mechanics*, 2 Vol., Springer, Berlin.
- K. Potthoff (1981), *Einführung in die Modelltheorie und ihre Anwendungen*, Wiss. Buchges., Darmstadt.
- E. Scheibe (1978), "On the structure of physical theories". In: *The Logic and Epistemology of Scientific Change*, Acta Phil. Fenn. 30, 2–4.
- E. Scheibe (1982), "Invariance and Covariance". In: J. Agassi, R. S. Cohen (eds.), *Scientific Philosophy Today. Essays in Honour of Mario Bunge*, Reidel, Dordrecht.
- E. Scheibe (1986), "Mathematics and Physical Axiomatization". In: *Mérites et Limites des Méthodes Logiques en Philosophie*, ed. by Fondation Singer-Polignac, Paris.
- H. J. Schmidt (1979), *Axiomatic Characterization of Physical Geometry*, Lecture Notes in Physics 111, Springer, Berlin.
- H. J. Schmidt (1981), "Stable Axioms in Physical Theories". In: A. Hartkämper, H. J. Schmidt (eds.), *Structure and Approximation in Physical Theories*, Plenum, New York.
- J. R. Shoenfield (1967), *Mathematical Logic*, Addison-Wesley, Reading.
- J. D. Sneed (1971), *The Logical Structure of Mathematical Physics*, Reidel, Dordrecht.
- P. Suppes (1957), *Introduction to Logic*, Van Nostrand, New York.

# Suppes Predicates for Classical Physics

NEWTON C. A. DA COSTA (São Paulo) / F. ANTONIO DORIA\* (Stanford)

## 1. Introduction

We have a threefold aim in the present paper: first, we wish to exhibit an unified treatment for the mathematical structures underlying what one usually calls in a loose way "classical physics", or "first-quantized physics", or even "classical field theory". That means, we are going to discuss Hamiltonian mechanics, electromagnetic theory in the Maxwell formulation, general relativity, the classical aspects of gauge field theory and the theory of the Dirac electron, again seen as a field theory. Such theories can also be looked upon as "first quantized theories" (but for Hamiltonian mechanics), as, for example, we can suppose that Einstein's gravitation is the depiction of the motion of a single graviton whose associated wave function is a nonlinear perturbation of a flat background metric field.

Now, since the *mathematical* structures we are going to deal with are set-theoretic constructs, we are going to develop them within a standard framework, such as the Zermelo-Fraenkel theory; that will be done with the help of Suppes predicates.

Finally, we wish to lay the groundwork for a systematic exploration of the consequences of metamathematical phenomena within theoretical and mathematical physics. We are especially interested in the consequences of, say, undecidability results that might appear within a given physical theory, or in the dependence of a given physical theory on a particular axiomatic system.

Section 2 of the present paper reviews the concept of Suppes predicate [39][40] in the da Costa-Chuaqui version [7]. Section 3 examines in detail our approach to the axiomatization of physical theories, which essentially follows two standard guidelines in the axiomatic treatments for physics, Klein's Erlangen Program [26] and the von Neumann formulation of quantum mechanics [23] [43]. Klein's ideas, as it is well-known, greatly influenced both the deve-

---

\* Partially supported by a Fulbright / CNPq-Brazil grant.

lopment of relativity theory and of modern field theories. In Section 4 we apply those concepts to obtain our formulation for "classical" and "first-quantized" theories out of a unified perspective. Finally Section 5 gives examples that, we hope, will show the usefulness of our treatment, while Section 6 sums up our ideas and evaluates our results.

The present paper is part of a series [7] [8] [9] [10] [11] [12] [13] dedicated to the exploration of the mathematical and philosophical foundations of physics in the light of modern metamathematical techniques and results.

## 2. *Suppes Predicates and Bourbaki Structures*

A review of the main set-theoretical concepts that we need here can be found in [11]. Our notation is standard; we follow [3] and [29].

Our main tool is the concept of Suppes predicate [39] [40] as a recipe for the axiomatization of physical theories; we follow a previous work [7] that has interpreted Suppes predicates as Bourbaki structures [4] [5]. The main idea goes as follows: Suppes notices [40] that we can try to directly axiomatize a theory by developing, say, a first-order language to handle the concepts in that theory. However, most concepts in an axiomatizable theory have usually been given a sound mathematical formulation. Now it is common mathematical practice to use some sort of informal set-theoretic language in the development of mathematical concepts. The Suppes predicate for a theory is simply the explicit construction of the concepts involved out of the set-theoretic background. The main advantage of Suppes' approach, besides the conceptual clarity it offers, is that we can easily move from informal, naïve-style set-theoretic discussions [37], to a rigorous axiomatic analysis in the style of foundational studies.

### *From Structures to Predicates*

A *mathematical structure*  $E$  is a finite ordered collection of sets (which may be particularized to relations and functions) of finite rank over the union of the ranges of two finite sequences of sets,  $X_1, X_2, \dots, X_m$  and  $Y_1, Y_2, \dots, Y_n$ , where  $m > 0$  and  $n \geq 0$ . If we are doing our constructions within ZFC,  $E$  is thus a ZFC set. The  $X$ 's and the  $Y$ 's are called the *base sets* of  $E$ ; the  $X$ 's are the *principal base sets*, while the  $Y$ 's are the *auxiliary base sets*.

The auxiliary base sets can be seen as previously defined structures, while the principal base sets are "bare" sets; for example, if we are describing a real vectorspace, the set of vectors is the only principal base set, while the set of scalars,  $\mathbb{R}$ , is the auxiliary base set; if we want to further specify things and talk about, say, 3-dimensional real Euclidean vector algebra, our principal base set will be given by the points in  $\mathbb{R}^3$ .

A *species of structures* or *Suppes predicate* is a formula of set theory whose only free variables are those explicitly shown:

$$P(E, X_1, X_2, \dots, X_m, Y_1, Y_2, \dots, Y_n),$$

such that  $P$  defines  $E$  as a mathematical structure on the principal base sets  $X_1, \dots, X_m$ , with the auxiliary base sets  $Y_1, \dots, Y_n$ , subject to restrictions imposed on  $E$  by the axioms we want our objects to obey. As the principal sets  $X_1, \dots$  vary over a class of sets in the set-theoretical universe, we get the structures of species  $P$ , or  $P$ -structures.

The Suppes predicate is then a conjunction of two parts: one specifies the set-theoretic process of construction of the  $P$ -structures, while the other imposes conditions that must be satisfied by the  $P$ -structures. This second conjunct contains the axioms for the species of structures  $P$ .

We can also write the Suppes predicate for  $E$  as follows:

$$Q(E) \leftrightarrow \exists X_1 \exists X_2 \dots \exists X_m P(E, X_1, X_2, \dots, X_m, Y_1, \dots, Y_n),$$

The auxiliary sets are thus seen as parameters in the definition of  $E$ ; the true "variables" are the principal base sets.

EXAMPLE 2.1 A group is an ordered quadruple

$$E = \langle X, f, g, e \rangle,$$

where  $X$  is a nonempty set,  $f$  is a binary operation on  $X$ ,  $g$  is a unary operation on  $X$ , and  $e \in X$  is a distinguished element. The group's elements are subject to the following conditions:

$$G1. \quad (x f y) f z = x f (y f z).$$

$$G2. \quad x f e = e f x = x.$$

$$G3. \quad x f x^g = x^g f x = e.$$

The corresponding Suppes predicate is:

$$P(E) \leftrightarrow \exists X \exists f \exists g \exists e (E = \langle X, f, g, e \rangle \wedge \phi(f, X) \wedge \psi(g, X) \wedge (e \in X) \wedge \forall x, y, z \in X (G1 \wedge G2 \wedge G3)).$$

where

$$\phi(f, X) \leftrightarrow "f \text{ is a function from } X \times X \text{ onto } X",$$

and

$$\psi(g, X) \leftrightarrow "g \text{ is a function from } X \text{ onto itself}."$$

Notice that in this example  $E$  has only one principal base set and no auxiliary base sets. We can easily obtain Suppes predicates for the usual algebraic structures, e.g. semigroups, rings, integral domains, vectorspaces, modules, linear, Lie and Grassmann algebras, and the like. If we consider, for example, vectors-

paces, we have a single principal base set (the set of vectors) and an auxiliary base set (the field of scalars).

Bourbaki notes [4] [5] that if we start with what he calls "mother structures", that is, algebraic, topological and ordered structures, all the remaining mathematical structures can be obtained out of combinations of the mother structures. We can thus define species of structures that encompass topological vectorspaces and their particularizations (Hilbert space, Fréchet space, LF-space, for example); Lie groups, differentiable manifolds, fiber bundles, and so on. All of everyday standard "professional" mathematics can be therefore axiomatized in a systematic way. Again following Bourbaki we can easily define isomorphism between structures, equivalent structures, initial structures, final structures and the like.

#### *Deduced and Derived Structures*

Given a structure  $E$  of species  $P(E, X_1, \dots, X_m, Y_1, \dots, Y_n)$ , let  $Z_1, \dots, Z_p$  be  $p$  ( $p > 0$ ) sets of finite rank over the union of ranges of the sequences

$$X_1, \dots, X_m, Y_1, \dots, Y_n ;$$

also let  $W_1, \dots, W_q$  ( $q \geq 0$ ) be  $q$  arbitrary sets. If the Suppes predicate

$$P^*(E^*, Z_1, \dots, Z_p, W_1, \dots, W_q)$$

defines  $E^*$  as a structure on the principal base sets  $Z_1, \dots$  with the  $W_1, \dots$  as auxiliary base sets, we say that the structure  $E^*$  of species  $P^*$  is *deduced* from the structure  $E$  of species  $P$ .

EXAMPLE 2.2 The species of structures of real vectorspaces has one principal base set (the set of vectors) and one auxiliary base set (the real scalars). From that species of structures we can deduce the underlying commutative group of vectors.

EXAMPLE 2.3 From a differentiable manifold  $M$  we can deduce at every point  $x \in M$  the local tangent space  $T_x$  and the tangent bundle  $TM$ .

We thus deduce "new" objects (local and global tangent bundles) out of the old one (the differentiable manifold  $M$ ).

We can obtain new structures out of (sets of) already defined structures by the means of two basic procedures:

1. With the help of set-theoretic operations, such as Cartesian products and passages to the quotient;
2. Through the imposition of new axioms to already-existing set-theoretic structures.

Therefore we can introduce the notion of *derived structure*. When we define a new structure  $E$  from a set  $S$  of other structures with the help of the two pro-



cedures described above, we say that  $E$  is *derived* from the structures  $S$ . The Suppes predicate of  $E$  can be expressed in terms of the Suppes predicates of the elements of  $S$ . We observe that the concept of *deduction* of structures is a particular case of derivation of structures.

The set  $S$  is the set of *ground structures* for  $E$ .

Finally, let  $E$  and  $E'$  be two structures of species  $P$  and  $P'$ , respectively. We suppose that  $P$  and  $P'$  differ only in connection with their sets of axioms, but that the conjunction of the axioms of  $P'$  implies each axiom of  $P$ , with quantifiers restricted to sets of finite rank over the union of the ranges of the base sets for  $E$ . If that is the case, we say that the  $P'$ -structure is *richer* than the  $P$ -structure (or that  $P'$  is richer than  $P$ ).

For instance, the species of commutative groups is richer than the species of groups.

The  $Q'$ -structure  $G'$  is then derived from the  $Q$ -structure  $G$  if  $Q'$  is richer than  $Q$ , or  $Q'$  can be obtained from  $Q$  in the way we have already described above.

**EXAMPLE 2.4** Consider a  $n$ -dimensional real differentiable manifold  $M$ ; let  $P(M, GL(n, \mathbb{R}))$  be the corresponding principal linear bundle, with  $GL(n, \mathbb{R})$  being the fiber and structure group. A reduction of that group to an orthogonal subgroup  $O(n, \mathbb{R})$ , possible whenever we endow  $M$  with a Riemannian differentiable metric tensor, leads to a species of structures (orthogonal principal bundles obtained as reductions of a principal linear bundle over  $M$ ) that is derived from the species of structures of principal linear bundles.

The above ideas can also be extended to the concept of *partial structures* introduced in [13].

### 3. Physical Theories

We follow traditional wisdom in seeing a physical theory as a triple

$$\mathcal{A} = \langle M, \Delta, \rho \rangle,$$

where (i)  $M$  is a Suppes-Bourbaki species of mathematical structures; (ii)  $\Delta$  is the theory's "domain of definition", and (iii)  $\rho$  gives the "interpretation rules" that relate  $M$  and  $\Delta$ . We can be more specific about (ii) and (iii), however, as we did elsewhere [9]; in any case we consider  $\mathcal{A}$  to be a set-theoretic construct.

**EXAMPLE 3.1** Let us consider the case of classical mechanics. We use it (say, in its Newtonian formulation) as a modelling tool in engineering, in astronautics and (partly) in semi-classical approximations as it is the case in plasma physics or in the "Berry phase" constructs. However there is a consensus that Newtonian mechanics cannot be universally applied; it has a restricted

domain  $\Delta$  of application. Very large objects (say, whole sections of the universe) are handled with another theory, Einstein's gravitation. And quantum mechanics is supposed to be valid in all known domains, but we restrict it to very small objects since its results agree with those of classical mechanics for objects of our size, modulo negligible deviations. Again in calculations that involve objects the size of the solar system, general relativity can be substituted by the classical models plus some perturbation theory. Thus, when we refer to a physical theory we must refer to its domain of application,  $\Delta$ . Moreover, while  $\Delta$  can be informally seen as a set-theoretic object, it pays to consider it within a standard axiomatic formulation for set theory [9] [14].

Now every physical theory encompasses a mathematical formalism  $M$ . Moreover all mathematics dealt with in physics can be easily fitted inside the framework of Zermelo-Fraenkel axiomatic set theory. Thus Hamiltonian mechanics can be seen as the theory of Hamiltonian flows on phase space, where phase space is an even-dimensional real symplectic manifold – that is to say, something that we can formalize in a pretty straightforward way inside axiomatic set theory. Thus there is a species  $M$  of mathematical structures called Hamiltonian mechanics, which we can investigate as any other mathematical species of structures.

Finally, both the domain  $\Delta$  and the Suppes predicate  $M$  aren't enough to determine  $\mathcal{A}$ . We must relate both, and the rules  $\rho$  we design for their interrelation are supposed to represent the way we use the mathematical constructs  $M$  within the concrete world; they are supposed to picture the "translation" between mathematics and reality. They are certainly the trickiest element in  $\mathcal{A}$ , since they embody the problem of the effectiveness of mathematics in depicting the world, but certainly several of its aspects can be clarified, as its logic [9], and tools inherited from statistics and the theory of measurement [28] [31] [39] [40]. Anyway, we assume that there is a system  $\rho$  of rules, explicitly or implicitly given, that determine the interconnection between our theory and its domain of application; those rules can sometimes be reduced to interpretation norms, in the sense of Stegmüller [35] [36] [37].

However we notice that if both  $M$  and  $\Delta$  are ZFC-sets, then for most usual physical theories  $\rho$  cannot be decidable [10]. We repeat our argument in Section 5 below, Remark 5.1.

**REMARK 3.1** Standard wisdom about the structure of theories tries to mimic inside  $M$  part of the facts and ideas about both  $\Delta$  and  $\rho$ . Thus in most cases  $M$  is built out of an ordered triple  $\langle S, O, L \rangle$ , where  $S$  is the set of states,  $O$  the observables in the theory, and  $L$  a measure space or some ordered structure (a propositional lattice) [25] [43] [23]. To go on with our example,  $S$  is phase space, the set of all points that can be "occupied" by a mechanical system.  $O$  is supposed to describe the set of all physically meaningful func-

tions defined on the states  $S$ ; every  $\sigma \in O$  is a map  $\sigma : K(S) \rightarrow L$ , where  $K(S)$  denotes some set constructed from  $S$ . In the case of mechanics one considers  $L$  to be a Boolean algebra of measurable sets in the real line  $\mathbb{R}$ , and the  $\sigma$  are maps from the Boolean algebra of some (or all) Borel sets in  $S$  onto  $L$ . Thus each observable is supposed to represent some kind of physical measurement process, since  $L$  can be given some (naïve) semantics as a set of questions of the form "the observable  $\sigma$  is in the set of states  $\Gamma \in S$ ".

We will deviate from the standard approach to the structure of  $M$ , since our main interest lies in making explicit the *mathematical* structures that underlie classical physics. In the conclusion, we will try to go back to the standard wisdom from our own point of view.

#### 4. Suppes Predicates for Classical Field Theories

In what follows we give a detailed analysis of some Suppes predicates for theories in the domain of what is usually called "classical physics", "classical field theory", or "first-quantized physics". We will try to exhibit their set-theoretic skeletons, and also what seems to be their main (formal, syntactical) unifying features.

From here on, for the sake of simplicity, when we talk about a physical theory we will always mean its Suppes predicate. That, of course, is an abuse of language.

The species of structures of essentially all physical theories can be formulated as particular dynamical systems derived from the triple  $P = \langle X, G, \mu \rangle$ , where  $X$  is a topological space,  $G$  is a topological group, and  $\mu$  is a measure on a set of finite rank over  $X \cup G$ . For example, a very general iteration process can be described out of a map of a compact space  $X$  onto itself; the process is the map's iteration. Its behavior can be characterized by the iteration's ergodic properties, while  $G$  can be taken to be the group of homeomorphisms of  $X$ . Any smooth dynamical system on a differentiable manifold  $X$  can be also derived from the properties of  $X$ ; one usually considers  $G$ -covariant objects, where now  $G$  is the group of diffeomorphisms of  $X$ .

The species of mathematical structures of physics, even in such a very general characterization, arises therefore out of geometry, out of the properties of a topological space  $X$ ; the physical objects are those that exhibit invariance properties with respect to the action of  $G$ . We are here following in a quite strict way the program sketched by Felix Klein, when he asserted that geometrical constructs should be obtained according to the following principle [26]:

*Given a manifold and a transformation group that acts on that manifold, develop the theory of the manifold's invariants with respect to the action of the group.*

However, we do not intend to delve into the most general situation. We will show in the sequel that the main species of structures in "classical" theories can be obtained out of two objects: a differentiable 'smooth' finite-dimensional real Hausdorff manifold  $M$  and a finite-dimensional Lie group  $G$ . Here 'smooth' means either 'of class  $C^k$ ',  $1 \leq k \leq +\infty$ , or, as it is increasingly frequent, that the functions defined on  $M$  are taken from some still more general function spaces, say, a Sobolev space or an arbitrary distribution space. In order to make those ideas more precise, we will examine in detail a particular example of a classical field theory: a well-established theory, namely, Maxwell's electromagnetism.

### *Maxwell's Electromagnetic Theory*

EXAMPLE 4.1 We start with Maxwell's theory as it might be presented by a mathematical physicist. Let  $M = \mathbb{R}^4$ , with its standard differentiable structure. Let us endow  $M$  with the Cartesian coordination induced from its product structure, and let  $\eta = \text{diag}(-1, +1, +1, +1)$  be the symmetric constant metric Minkowskian tensor on  $M$ . If the  $F_{\mu\nu}(x)$  are components of a 'smooth' covariant 2-tensor field on  $M$ ,  $\mu, \nu = 0, 1, 2, 3$ , then Maxwell's equations are:

$$\begin{aligned}\partial_\mu F^{\mu\nu} &= j^\nu, \\ \partial_\mu F_{\nu\rho} + \partial_\rho F_{\mu\nu} + \partial_\nu F_{\rho\mu} &= 0.\end{aligned}$$

The contravariant vectorfield whose components are given by the set of four 'smooth' functions  $j^\mu(x)$  on  $M$  is the current that serves as source for Maxwell's field  $F_{\mu\nu}$ . Here by 'smooth' we might even mean a piecewise differentiable function, to account for shock-wave like solutions.

It is known that Maxwell's equations are equivalent to the Dirac-like set

$$\nabla\varphi = \iota,$$

where

$$\varphi = (1/2)F_{\mu\nu}\gamma^{\mu\nu},$$

and

$$\begin{aligned}\iota &= j_\mu\gamma^\mu, \\ \nabla &= \gamma^\rho\partial_\rho,\end{aligned}$$

(where the  $\{\gamma^\mu : \mu = 0, 1, 2, 3\}$  are the Dirac gamma matrices with respect to  $\eta$ ). Those equation systems are to be understood together with boundary conditions that specify a particular field tensor  $F_{\mu\nu}$  "out of" the source  $j^\nu$ . For a reference see [19].

The symmetry group of the Maxwell field equations is the Lorentz-Poincaré group that acts upon Minkowski space  $M$  and in an induced way on objects defined over  $M$ . However, since we are interested in *complex* solutions for the Maxwell system, we must find a reasonable way of introducing complex objects in our formulation.

The way one usually does it is to formalize the Maxwellian system as a gauge field. We now sketch the usual formulation: again we start from the manifold  $M = \langle \mathbb{R}^4, \eta \rangle$ , which is Minkowski spacetime, and construct the trivial circle bundle  $P = M \times S^1$  over  $M$ , since Maxwell's field is the gauge field of the circle group  $S^1$  (usually written in that respect as  $U(1)$ ). We then form the set  $\mathcal{E}$  of bundles associated to  $P$  whose fibers are finite-dimensional vectorspaces. The set of physical fields in our theory is obtained out of some of the bundles in  $\mathcal{E}$ : the set of electromagnetic field tensors is a set of cross-sections of the bundle  $F = \Lambda^2 \otimes s^1(M)$  of all  $s^1$ -valued 2-forms on  $M$ , where  $s^1$  is the group's Lie algebra. To be more precise, the set of all electromagnetic fields is a manifold  $\mathcal{F} \subset C^k(F)$ , if we are dealing with  $C^k$  cross-sections (it is a submanifold in the usual  $C^k$  topology due to the closure condition  $dF = 0$ ).

Finally we have a group action on  $\mathcal{F}$ : in fact, two groups act on the electromagnetic fields. The first one is the Lorentz-Poincaré group  $L$  that will be here seen as a subgroup of the group of diffeomorphisms of  $M$ ; the second group action it is the (in the present case trivial) action of the group  $\mathcal{G}'$  of gauge transformations of  $P$  when acting on the field manifold  $\mathcal{F}$ . As it is well known, its action is *not* trivial in the non-Abelian case. Also it has a non-trivial action on the space  $\mathcal{A}$  of all gauge potentials of the fields in  $\mathcal{F}$ . Therefore we take as our symmetry group  $\mathcal{G}$  the product  $L \otimes \mathcal{G}'$  of the (allowed) symmetries of  $M$  and the symmetries of the principal bundle  $P$ . For mathematical details see [20].

We must also add the spaces of potentials,  $\mathcal{A}$ , and of currents,  $I$ , as structures derived from  $M$  and  $S^1$ . Both spaces have the same underlying topological structure; they differ in the way the group  $\mathcal{G}'$  of gauge transformations acts upon them. We then obtain  $I = \Lambda^1 \otimes s^1(M)$  and  $\mathcal{A} = I = C^k(I)$ . However notice that  $I/\mathcal{G}' = I$ , since  $\mathcal{G}'$  acts in a trivial way on  $I$ , while  $\mathcal{A}/\mathcal{G}' \neq \mathcal{A}$ .

Therefore we can say that the 9-tuple

$$\langle M, S^1, P, \mathcal{F}, \mathcal{A}, \mathcal{G}, I, B, \nabla\varphi = \mathfrak{t} \rangle$$

where  $M$  is Minkowski space, and  $B$  is a set of boundary conditions for our field equations  $\nabla\varphi = \mathfrak{t}$ , represents the species of mathematical structures of a Maxwellian electromagnetic field, where  $P$ ,  $\mathcal{F}$  and  $\mathcal{G}$  are derived from  $M$  and  $S^1$ , in the sense discussed before Example 2.4. The Dirac-like equation

$$\nabla\varphi = \mathfrak{t}$$

should be seen as an axiomatic restriction on our objects; the boundary conditions  $B$  are (i) a set of derived species of structures from  $M$  and  $S^1$ , since, as we are dealing with Cauchy conditions, we must specify a local or global spacelike hypersurface  $C$  in  $M$  to which (ii) we add sentences of the form  $\forall x \in C \theta(x) = \theta_0(x)$ , where  $\theta_0$  is a set of (fixed) functions and the  $\theta$  are the adequate restrictions of the field functions and equations to  $C$ .

To make things clear, we may split that 9-tuple into the following pieces:

1. *The Ground Structures.* These are the couple  $\langle M, S^1 \rangle$ .
2. *The Derived Field Spaces.* They are the potential, field and current spaces  $\langle \mathcal{A}, \mathcal{F}, I \rangle$ .
3. *The Symmetry Group.* In the present case, it is  $\mathcal{G} = L \otimes \mathcal{G}'$ .
4. *An Axiom for the Dynamics of the System.* It is given by the Dirac-like equation  $\nabla\varphi = \iota$ , together with the boundary conditions  $B$ .
5. *Intermediate Sets.* Sets that appear in our construction but that do not have a direct physical meaning; it is the case of the principal bundle  $P$  and associated tensor bundles.

We will extend that recipe to encompass all other examples of classical field theories, as listed in the beginning of this paper.

### *Classical Physics*

The preceding example allows us to try a general characterization for classical field theories:

**DEFINITION 4.1** *The species of structures of a classical physical theory is given by the 9-tuple*

$$\Sigma = \langle M, G, P, \mathcal{F}, \mathcal{A}, I, \mathcal{G}, B, \nabla\varphi = \iota \rangle,$$

*which is thus described:*

1. *The Ground Structures.* They are the couple  $\langle M, G \rangle$ , where  $M$  is a finite-dimensional "smooth" real manifold, and  $G$  is a finite-dimensional Lie group.
2. *The Intermediate Sets.* A fixed principal fiber bundle  $P(M, G)$  over  $M$  with  $G$  as its fiber, as well as several associated tensor and exterior bundles.
3. *The Derived Field Spaces.* They are the potential space  $\mathcal{A}$ , the field space  $\mathcal{F}$  and the current or source space  $I$ .  $\mathcal{A}$ ,  $\mathcal{F}$  and  $I$  are spaces (in general, manifolds) of cross-sections of the bundles that appear as intermediate sets in our construction.
4. *Axiomatic Restrictions on the Fields.* They are given by the dynamical rule  $\nabla\varphi = \iota$  and by the relation  $\varphi = d(\alpha)\alpha$  between a field  $\varphi \in \mathcal{F}$  and its

potential  $\alpha \in \mathcal{A}$ , together with the corresponding boundary conditions  $B$ . Here  $d(\alpha)$  denotes a covariant exterior derivative with respect to the connection form  $\alpha$ , and  $\nabla$  is a covariant Dirac-like operator.

5. The Symmetry Group.  $\mathcal{G} \subseteq \text{Diff}(M) \otimes \mathcal{G}'$ , where  $\text{Diff}(M)$  is the group of diffeomorphisms of  $M$  and  $\mathcal{G}'$  is the group of gauge transformations of the principal bundle  $P$ .
6. The Space of Physically Distinguishable Fields. If  $\mathcal{K}$  is one of the  $\mathcal{F}$ ,  $\mathcal{A}$  or  $\mathcal{I}$  field manifolds, then the space of physically distinct fields is
 
$$\mathcal{K}/\mathcal{G},$$

according to Klein's prescription.

It remains to show that what we call "classical physics" fits easily into that formulation; in particular we must show that their dynamics are given by a Dirac-like equation. That will be ascertained when we deal with each specific situation.

#### *The Construction of Manifolds and Lie Groups*

The mathematical species of structures that characterize a differentiable finite-dimensional real manifold  $M$  can be built from one principal base set  $X$  and one auxiliary set,  $\mathbb{R}$ .  $X$  is supposed to be a separable complete metric space to which we will impose further restrictions through the Suppes predicate that defines the manifold.

We then proceed to define several sets of finite rank over  $X \cup \mathbb{R}$  that will be needed in the sequel. First we get  $\mathbb{R}^n$ ,  $n < \omega_0$ , where  $\omega_0$  is the set of natural numbers;  $n$  is kept fixed and is the manifold's dimension (the Cartesian product is a finite-rank operation). We then form the infinite product sets  $\mathbb{R}^X$  and  $(\mathbb{R}^n)^X$ . If we are dealing with, say,  $C^k$  objects, for  $1 \leq k \leq +\infty$ , we obtain the subsets

$$C^k(X, \mathbb{R}) \subset \mathbb{R}^X,$$

and

$$C^k(X, \mathbb{R}^n) \subset (\mathbb{R}^n)^X.$$

To get the product sets  $\mathbb{R}^X$  and  $(\mathbb{R}^n)^X$  we have used the power set axiom; the subsets we have described above can be obtained from the corresponding supersets with the help of the separation axiom. We will also need restrictions like  $C^k(U, \mathbb{R})$  and  $C^k(U, \mathbb{R}^n)$ , where  $U \subset X$  is an open set; again one uses the separation axiom as our tool in obtaining the restriction. Finally, in order to define a differentiable manifold via local coordinations we will also require  $C^k(X, X)$  and  $C^k(\mathbb{R}^n, \mathbb{R}^n)$ . Derived species of structures like tangent spaces at a point, the tangent bundle, the dual cotangent bundle; tensor, cotensor,

symmetric and exterior bundles, can be obtained by following the usual procedures in differential geometry [38].

Finally, a Lie group is a group whose underlying set has the structure of a differentiable manifold, and such that the group's product operation is differentiable with respect to the manifold's structure. We thus derive the mathematical species of structures of a Lie group out of those for a group and for a differentiable manifold.

### *General Relativity*

General relativity is a theory of gravitation that relates gravitational forces to the (pseudo)metric structure of spacetime through the Einstein equations. Given any 4-dimensional noncompact real "smooth" manifold  $M$ , we can endow it with a continuum of nonequivalent (that is, nondiffeomorphic) Lorentzian metric tensors. Therefore, neither the underlying structure of  $M$  as a topological manifold nor its (sometimes quite complicated) differentiable structure [11][42] determines its pseudo-Riemannian metric tensor. From the strictly geometrical viewpoint, when we choose a particular Lorentzian metric tensor  $g$ , we determine a reduction of the general linear bundle over  $M$  to one of its (differently embedded) pseudo-orthogonal subbundles. The relation is 1-1. We then follow our recipe:

- We take as ground structures a 4-dimensional real "smooth" manifold  $M$ , and the Lorentz pseudo-orthogonal group  $O(3,1)$
- We then form the principal linear bundle  $L(M)$  over  $M$ ; that structure is solely derived from  $M$ , as it arises from the covariance properties of the tangent bundle over  $M$ . From  $L(M)$  we fix a reduction of the bundle group,  $L(M) \rightarrow P(M, O(3,1))$ . Those will be our intermediate sets. We therefore define a Lorentzian metric tensor  $g$  on  $M$ , and get the spacetime  $\langle M, g \rangle$ .
- Therefore the general-relativistic spacetime arises quite spontaneously out of the interplay between the theory's "general covariance" aspects (which appear in the linear bundle  $L(M)$ ) and its "gauge-theoretic" features (which give rise to the principal bundle  $P$ ).
- Now for the field spaces. We start with three different field spaces. The first is the set (manifold, if we are dealing with a  $C^k$  set of objects) of all metric tensors  $\mathcal{M} \subset C^k(\otimes^2 T \cdot (M))$  on  $M$ , where  $\otimes^2 T \cdot (M)$  is the bundle of all symmetric real-valued 2-forms on  $M$ . Out of  $\mathcal{M}$  we get  $\mathcal{A}$ , which is the space of all Christoffel connections on  $M$ , and  $\mathcal{F}$ , the space of all Riemann-Christoffel curvature tensors on  $M$ .

We deal with two source fields. One of them,  $I'$ , is the space of all momentum-energy tensors, and gives the source for the Einstein equations. The other source space is derived from  $I'$ , and is a space  $I$  of higher-order tensors which gives the source for the Dirac-like dynamics for our objects.



- $\mathcal{G}$  is the group of  $C^k$ -diffeomorphisms of  $M$ .
- If  $\mathcal{K}$  is any of the field spaces above, the space of physically distinct fields is the quotient  $\mathcal{K}/\mathcal{G}$ .
- Finally, the dynamics in our picture for general relativity. The Dirac-like equation for general relativity is equivalent to the Einstein equations, given adequate boundary conditions [18] [19] (the last reference discusses linearized gravitation).

### *Classical Gauge Fields*

The mathematics of classical gauge fields can be found in [2] [6] [27]. We follow here the preceding examples:

- The ground structures are a spacetime  $\langle M, g \rangle$  and a finite-dimensional semi-simple Lie group  $G$ .
- The intermediate set is a fixed principal bundle  $P(M, G)$  over  $M$  with  $G$  as its fiber.
- We then get connection or potential space  $\mathcal{A}$ , which coincides with the space of all  $C^k$  cross-sections of the bundle of  $\ell(G)$ -valued 1-forms on  $M$ , where  $\ell(G)$  is the group's Lie algebra. Curvature, or field space  $\mathcal{F}$  is the space of all  $C^k$  cross-sections of  $\ell(G)$ -valued 2-forms on  $M$  such that, for  $\varphi \in \mathcal{F}$  and  $\alpha \in \mathcal{A}$ ,  $\varphi = d(\alpha)\alpha$ . Source space  $I$  coincides with  $\mathcal{A}$ , but is acted upon in a different way by the group  $\mathcal{G}$  of gauge transformations, since "currents" are *tensorial* 1-forms, while "gauge potentials" are *pseudo-tensorial* 1-forms.
- The space of physically different fields is  $\mathcal{K}/\mathcal{G}$ , where  $\mathcal{K}$  denotes any of the above field spaces.
- The gauge field equations can be formulated as a Dirac-like equation [21], and suitable boundary conditions  $B$  can make that Dirac-like equation fully equivalent to the usual field equations. Or we can add an extra condition to the Dirac-like equation. For the gauge field equations are

$$\delta(\alpha)\varphi = \mathfrak{t},$$

and

$$d(\alpha)\varphi = 0$$

$\delta$  is the covariant divergence (dualized from  $d$ ) with respect to the connection form  $\alpha$ ;  $\mathfrak{t}$  is the source for the gauge field. The second equation is the Bianchi differential identity. We can add that

$$\varphi = d(\alpha)\alpha,$$

since noncurvature fields may satisfy the differential Bianchi conditions [22] for degenerate  $\alpha$ . Now if we write

$$\nabla(\alpha) = d(\alpha) - \delta(\alpha),$$

we can write the gauge field equations in the Dirac-like form as

$$\nabla(\alpha)\varphi = \mathbf{1}.$$

To ensure unicity we must add the extra condition  $\varphi = d(\alpha)\alpha$ .

### *Unified Theories and Classical Superfields*

The extension of the preceding constructions to Kaluza-Klein-like unified theories (at the classical level) is straightforward: see [6]. If we slightly modify the construction of a differentiable manifold's species of structures – if we add a new auxiliary base set, a Grassmann algebra – we can include in the present scheme the classical theory of supermanifolds and their superfields; for a reference [17].

### *The Dirac Electron on a Spacetime Manifold*

We look at the Dirac electron as a classical field coupled to another classical field, Einstein's gravitation, and not from the more familiar, quantum-mechanical perspective.

Given a spacetime  $\langle M, g \rangle$ , we can deduce the structure of a Clifford bundle out of the tangent and cotangent bundles  $T \cdot M$  and  $T^* \cdot M$ . If the manifold  $M$  admits a spinor structure, we can form a spinor bundle over  $M$ ; its sections are spinor fields, and the Dirac electron on a spacetime manifold satisfies Dirac's equation coupled to the gravitational field through a tetrad field. If  $M$  is Minkowski space, we get the usual Dirac theory.

We can also obtain the Schrödinger equation as the nonrelativistic flat-space limit of our construction for the Dirac electron.

### *Hamiltonian Mechanics*

Hamiltonian mechanics is seen as the dynamics of the "Hamiltonian fluid" [1], that is, again as a field theory. Our ground species of structures are a  $2n$ -dimensional real smooth manifold, and the real symplectic group  $\text{Sp}(2n, \mathbb{R})$ . Phase spaces in Hamiltonian mechanics are symplectic manifolds, that is, even-dimensional manifolds like  $M$  endowed with a symplectic form, that is, a nondegenerate closed 2-form  $\Omega$  on  $M$ . As in the case of general relativity, the imposition of that form can be seen as the choice of a reduction of the linear bundle  $L(M)$  to a fixed principal bundle  $P(M, \text{Sp}(2n, \mathbb{R}))$ ; however in the present case given one such reduction it doesn't automatically follow that the induced 2-form on  $M$  is a closed form.

All other objects are constructed in about the same way as in the preceding examples. However we must show that we still have here a Dirac-like equation as the dynamical axiom for the species of structures of mechanics. Hamilton's equations are

$$i_X \Omega = -dh,$$

where  $i_X$  denotes the interior product with respect to the vectorfield  $X$  over  $M$ , and  $h$  is the Hamiltonian function. That equation is (locally, at least) equivalent to:

$$L_X \Omega = 0,$$

or

$$d(i_X \Omega) = 0,$$

where  $L_X$  is the Lie derivative with respect to  $X$ . The condition  $d\varphi = 0$ , with  $\varphi = i_X \Omega$ , is the degenerate Dirac-like equation for Hamiltonian mechanics.

We don't get a full Dirac-like operator  $\nabla \neq d$  because  $M$ , seen as a symplectic manifold, doesn't have a canonical metrical structure, so that we cannot define (through the Hodge dual) a canonical divergence  $\delta$  dual to  $d$ .

The group that acts on  $M$  with its symplectic form is the group of canonical transformations; it is a subgroup of the group of diffeomorphisms of  $M$  so that symplectic forms are mapped onto symplectic forms under a canonical transformation. We can take as "potential space" the space of all Hamiltonians on  $M$  (which is a rather trivial function space), and as "field space" the space of all "Hamiltonian fields" of the form  $i_X \Omega$ .

#### *Two Comments*

**REMARK 4.1** We have thus derived the main species of structures in classical physics from two ground structures, a finite-dimensional smooth real manifold  $M$  and a semi-simple finite-dimensional Lie group  $G$ . However, given an arbitrary pair  $\langle M, G \rangle$ , it obviously doesn't follow that such a pair represents the ground structures for some theory in classical physics. Therefore it seems that we still need some kind of "superselection rule" that would separate the pairs  $\langle M, G \rangle$  which give rise to the mathematics of classical physics from those without any clear physical meaning.

We envisage two possibilities at this point: either Mother Nature organizes the universe at the "classical physics" level out of couples like  $\langle M, G \rangle$ , plus some hidden superselection criterion, which however we may expect to formulate in a rigorous and simple way sometime in the future, or we have somehow restricted our way of perceiving the universe so that we can only "understand" it at the classical level out of manifolds and symmetries that act upon them.

That second alternative, with its almost Kantian flavor, is certainly much more complex than the first, and we see no way it could be ascertained in a secure way. We will end up either by discussing the "cultural background" behind Klein's Program, or looking for some theory of knowledge that might

justify it, and will be led astray from our main point. However its complexity looks like a sign of its truthfulness.

We will elaborate on that dilemma elsewhere. We might summarize it as a discussion between those that think that the universe "is" that way, and those that believe that we have been "led to believe that the universe is that way".

REMARK 4.2 There is a formal advantage in our presentation of the structures for classical physics: we can easily reformulate them in the language of K-theory: classical physics is a subcategory of the category of  $G$ -equivariant smooth vectorbundles over a smooth real manifold  $M$ . We thus start to bridge the gap between axiomatic formulations for classical physics and the category treatment for the same subject. Also, due to the crossroads nature of K-theory, we are allowed several interesting and even unexpected perspectives to envisage the mathematics of classical physics in a global, unified way.

## 5. Examples

We axiomatize a theory in order to obtain metatheorems about that theory. We are therefore interested in, say, proving that a given axiomatic treatment for a physical theory is undecidable, or incomplete, or to obtain examples (if any) of physically meaningful undecidable statements within that theory.

Also we may be interested in model-theoretic constructions: as we have embedded classical physics inside ZFC, we can examine the manifold models for that set theory looking for model-theoretic phenomena that might be given some physical meaning.

It is obvious that the crucial idea here is the (rather loose) concept of "physically meaningful phenomenon". We won't try to define such a concept. However we will presume (or at least hope) that our main examples, which deal with objects defined within physical theories, and that discuss problems formulated within the usual intuitive mathematical constructions of physics, will somehow satisfy that criterion.

### *Undecidability and Incompleteness in Classical Mechanics*

We state here a general undecidability and incompleteness result that has been discussed and proved in [10] and apply it to classical mechanics.

We need a well-known result: let  $L_{Ar}^1$  be the first-order language of axiomatized arithmetic [34]. Let  $N$  be the standard model for axiomatized arithmetic.

Let  $T$  be a first-order theory whose language  $L_T \supset L_{Ar}^1$ . When we explicitly say that  $T$  is consistent, this will mean that any arithmetical theorem of  $T$  is true in the standard model  $N$  of arithmetic. Therefore such a condition implies that  $T$  is consistent in the usual sense, that is, it doesn't contain contradictory theorems.

We thus consider a first-order theory  $T$  such that  $L_T^1 \supset L_{Ar}^1$ . Suppose that the set of axioms of  $T$  is recursively enumerable. Suppose also that one can prove from  $T$  every statement of the form  $m+n=p$ ,  $m.n=p$  and  $m < n$ , where  $m, n \in \omega_0$ , which is true in the standard model  $N$ . Then:

**PROPOSITION 5.1** *We can construct in  $T$  a Diophantine equation*

$$p(x_1, \dots, x_n) = 0$$

*so that  $p = 0$  has no solutions in the natural numbers in  $N$ , but such that*

$$T \not\vdash \neg (\exists \vec{x} \in \omega_0^n p(\vec{x}) = 0),$$

*where we abbreviate  $\vec{x} = \langle x_1, \dots, x_n \rangle$ .*

*Proof:* See [15] [16]. Notice that we also have

$$T \not\vdash (\exists \vec{x} \in \omega_0^n p(\vec{x}) = 0)$$

since in that case all the arithmetical theorems of  $T$  wouldn't be true in  $N$ . Therefore the sentence

$$\exists \vec{x} \in \omega_0^n p(\vec{x}) = 0$$

is undecidable in  $T$ , for this particular  $p$ .

**COROLLARY 5.1** *If ZFC is consistent, then neither  $ZFC \vdash \exists \vec{x} \in \omega_0^n p(\vec{x}) = 0$  nor  $ZFC \vdash \neg (\exists \vec{x} \in \omega_0^n p(\vec{x}) = 0)$ . Moreover, for a model  $M$  of ZFC such that  $N \subset M$  is a model for arithmetic in the theory, then  $M \models \neg (\exists \vec{x} \in \omega_0^n p(\vec{x}) = 0)$*

The condition on  $N$  means that we exclude nonstandard models for arithmetic within our theory.

From that result we can prove a general incompleteness theorem for elementary functions of a real variable. The algebra  $\mathcal{A}$  of elementary functions of a real variable includes the polynomials,  $\sin x$ ,  $\log x$ ,  $e^x$ , the constant  $\pi$ , and is closed under finite sums, products, multiplication by rational numbers, and function composition. We mainly rely on Richardson's functor, that translates results about Diophantine equations into results about elementary functions of a real variable [10] [33]

As we noticed, there is a general undecidability and incompleteness result at work here [41]. Everything proceeds within ZFC (or within any similarly powerful axiomatic system), so that we can obtain all the maps given by Richardson's functor into  $\mathcal{A}$  and extensions. Let  $Q[Q]$  be the (denumerable) algebra of rational functions  $p/q$  over  $Q$ , where  $p, q \in \mathcal{A}$ , and let  $\mathcal{B} \supseteq Q$  be a denumerable superalgebra that includes  $Q$ . Let  $\psi \in L_{ZFC}^1$  be a predicate defined for  $\mathcal{B}$  such that we can effectively obtain  $f_0, g_0 \in \mathcal{B}$  and  $ZFC \vdash \psi(f_0)$  so that  $ZFC \vdash \neg \psi(g_0)$ . Then:

PROPOSITION 5.2 *If ZFC is consistent, then:*

1. *There is an  $h \in \mathcal{B}$  so that neither  $\text{ZFC} \vdash \neg\psi(h)$  nor  $\text{ZFC} \vdash \psi(h)$ , but  $M \models \psi(h)$ .*
2. *There is a denumerable set of functions  $h_m(x) \in \mathcal{B}$  such that there is no general decision procedure to ascertain, for an arbitrary  $m$ , whether  $\psi(h_m)$  or  $\neg\psi(h_m)$ .*

As a corollary to those results we can state several undecidability and incompleteness results in Hamiltonian mechanics: let  $P$  be a phase space. Then we have the following:

PROPOSITION 5.3 *In ZFC, we have that:*

1. *There is no algorithm to check, for an arbitrary Hamiltonian  $h$  and a smooth function  $f$  on  $P$ , whether  $f$  is a first integral of  $h$ .*
2. *For an arbitrary Hamiltonian  $h$  whose associated Hamiltonian dynamical system has been proved to be integrable by quadratures, there is no general algorithm to solve the corresponding Hamilton-Jacobi equation.*
3. *There is a denumerable set  $h_k, k \in \omega_0$ , of Hamiltonians defined on an open starshaped domain  $U \subseteq P$  so that there is no general algorithm to check, for arbitrary  $k$ , whether the associated Hamiltonian systems  $X_{h_k}$  can be integrated by quadratures.*

*Proof:* See [10].

Some of these results on algorithmic impossibility can also be extended to incompleteness results. We remember that all our models for ZFC include the standard model  $N$  for arithmetic.

PROPOSITION 5.4 *If ZFC is consistent, then there is a Hamiltonian system of which it is true (of a model  $M$  for ZFC) that it cannot be integrated by quadratures, but such that this fact cannot be proved in the given axiomatization for symplectic geometry.*

*Proof:* A formal proof can be directly obtained out of Proposition 5.2.

However we can offer here an informal argument that still has the flavor of a proof that goes back to Post in 1944.

We imitate [15]. We restrict our attention to a denumerable algebra  $\mathcal{B} \supset \mathcal{Q}$  [Q]. We generate all the theorems in the given axiomatization. Within such a listing we form two sublists: list A contains the Hamiltonians whose associated vectorfields can be provably integrated by quadratures. List A is recursively enumerable, we know, since we can work it backwards from the elements of  $\mathcal{B}$ , which is a countable set. List B contains those that we have proved that *cannot* be integrated by quadratures. Now list B cannot contain all Hamiltonian systems that cannot be thus integrated. For if it did, the set of all those

Hamiltonians which cannot be "nicely" integrated would be recursively enumerable, and there would be a decision procedure for integrability. Thus there is a Hamiltonian in our theory that cannot be integrated by quadratures, but that fact cannot be proved within the given axiomatization.

### *Incompleteness of Chaos*

Can we check whether there is chaos in a dynamical system? Given the convoluted trajectory of a computer-integrated vectorfield, can we be sure that sheer mathematical ingenuity will someday allow us to show that any such system that looks chaotic is, in fact, chaotic according to some sound mathematical criterion? No:

PROPOSITION 5.5 *If ZFC is consistent, then:*

1. *There is an ergodic dynamical system on  $\mathbb{R}^2$  so that its ergodicity cannot be proved from the axioms of the theory.*
2. *There is a dynamical system with a Smale horseshoe, but such that the existence of the horseshoe cannot be proved from the axioms of ZFC.*
3. *There is a Bernouillian flow that cannot also be proved to be so.*

*Proof:* Follows directly from 5.3. See [10].

### *Set-Theoretic Genericity and General Relativity*

We have recently investigated the relevance of the concept of set-theoretically genericity in the realm of general relativity [11]. The main question lurking behind our work was: can we *physically* detect set-theoretic genericity in the world around us? Again we deal with an imprecise concept, the idea of "physically detecting" some mathematical property. However, as we now show, our work has amounted to showing that some systems in general relativity may be generic, while others aren't definitively generic in the set-theoretic sense.

We first dealt with cylindrical spacetimes, that is, those spacetimes homeomorphic to  $C \times \mathbb{R}$ , where  $C$  is a compact smooth 3-manifold. We then proved in [11]:

PROPOSITION 5.6 *Every cylindrical spacetime is standard.*

This means that forcing extensions do not add new objects to the set of cylindrical spacetimes. However, if  $V$  is our standard set-theoretic universe, and if  $V(\mathcal{G}) \supset V$  is a forcing (or a Boolean) extension of the ZFC universe  $V$  such that there are set-theoretically generic real numbers in  $V(\mathcal{G})$ , then:

PROPOSITION 5.7  $V(\mathcal{G}) \models$  "There is a set-theoretically generic spacetime".

Moreover, we have a stricter result:

**PROPOSITION 5.8**  $V(g) \models$  "*There is a set-theoretically generic spacetime homeomorphic to  $\mathbb{R}^4$* ".

To summarize: if we live in a cylindrical spacetime, then our universe isn't generic; and even if our universe is topologically flat, it can be set-theoretically generic.

Related results are:

**PROPOSITION 5.9**

1. *Open balls in a spacetime are diffeomorphic to a standard open domain.*
2. *Compact domains with a smooth boundary in spacetime are standard.*

*Comment:* This means that laboratories are in standard domains, and that all subluminal information that one gets in the course of a physical observation of the universe (from the point of view of general relativity) is included in a standard domain.

We have suggested that set-theoretic genericity could be intuitively assimilated to randomness. However we proved that such an identification *cannot* be verified within ZFC, that is to say, it is independent of the axioms of set theory:

**PROPOSITION 5.10** *The sentence "Every set-theoretically generic spacetime is random in the sense of Kolmogorov-Chaitin-Martin-Löf, modulo a meager set of spacetimes" is independent of ZFC.*

*Comment:* Notice that the space of all spacetimes  $\mathcal{M} \subset \mathcal{P}(\mathbb{R}^9)$  due to the Whitney embedding theorem, where  $\mathcal{P}$  denotes the power set. Therefore  $\mathcal{M}$  is a ZFC set. In the preceding proposition, we code noncompact spacetimes by a set of binary irrational reals in the unit interval  $[0,1]$ , modulo homeomorphisms of that interval. Therefore we can argue about the set of generic and the set of KCML-random binary irrationals in a pretty straightforward way.

Our exploration of genericity in gravitation theory was just a preliminary effort, but we think that the examples we have given suggest that there is much more in stock when a deeper and more systematic investigation is pursued.

#### *Related Results of Some Interest*

We can also quote another simple example of an undecidable sentence within ZFC which deals with physical concepts. We say that a map is a "topological isomorphism" when it is an algebraic isomorphism that induces an homeomorphism of the underlying topological spaces; also, if  $X$  and  $\alpha$  are sets,  $X^\alpha$  is a product space.

**PROPOSITION 5.11** *The sentence below within quotation marks is*



*undecidable with respect to the ZFC axioms:*

*Let  $S^1 = \mathbb{R} / \mathbb{Z}$ , with the quotient topology. Then: "If  $\models (S^1)^\alpha \models 2^{\aleph_0}$ , then  $(S^1)^\alpha$ , with the product topology, is topologically isomorphic to a Lie group".*

*Comment:* Proof is a simple consequence of the undecidability of  $2^{\aleph_1} > 2^{\aleph_0}$ .

We give a first consequence of that result:

**PROPOSITION 5.12** *The sentence "Every compact arcwise-connected topological group of the cardinality of the continuum is topologically isomorphic to the space of action-angle variables of a denumerable system of independent harmonic oscillators" is independent of the ZFC axioms.*

*Comment:* The action-angle variable space is, in our case, at most a countable torus  $(S^1)^{\omega_0}$ , which is a Lie group, and also a compact, arcwise-connected topological group of cardinality  $2^{\aleph_0}$ . The sentence is true in the constructive universe  $L$ , and false in a model that satisfies Martin's axiom.

While such a sentence isn't as striking as the other examples of undecidable assertions that we have offered, we believe one can obtain some interesting results along these paths, since we are actually dealing here with very simple (and yet undecidable) assertions in point-set topology.

#### *The Map $\rho$ between $M$ and $\Delta$ Cannot be Decidable*

**REMARK 5.1** Notice that our results imply that, if a physical theory is given by the tripartite structure  $\Phi = \langle M, \Delta, \rho \rangle$ , then in nontrivial situations  $\rho$  cannot be decidable, if  $M$  and  $\Delta$  are seen as sets within a theory like ZFC. For we can generate out of Proposition 5.2 a family of Hamiltonians  $h_k$ ,  $k \in \omega_0$ , such that each  $h_k$  is, say, either a free particle or a harmonic oscillator, but there is no general algorithmic procedure to determine, for an arbitrary  $k$ , which is which. Therefore  $\rho(h_k) \in \Delta$  is either a free particle or a harmonic oscillator, but we cannot in general determine which option is valid, since then we would have an effective procedure to decide  $\{h_k\}$ , and we have shown that there is no such a decision procedure.

## *6. Conclusions and Acknowledgments*

The main features of our work can be thus summarized:

- We have shown that what one usually calls "classical physics" or "classical field theory" can be axiomatically formulated with the help of Suppes predicates [7] as species of structures derived from a couple  $\langle M, G \rangle$ , where  $M$  is a smooth real finite-dimensional differentiable manifold, and  $G$  is a semi-simple, finite-dimensional Lie group.

- Out of that axiomatization we obtain undecidability and incompleteness results in classical mechanics and in general relativity; it is however clear that those results can be easily extended to other similarly axiomatized field theories.
- We have also shown that those undecidability results extend to the mainstream conception of a scientific theory as a tripartite structure – a mathematical model, a domain of interpretation, and a rule that connects the mathematics to its interpretation. If we have an actual mathematical model for a physical theory, then the connection between mathematics and its interpretation cannot be given by an effective rule (if we suppose that the domain of interpretation can be embedded into an adequately axiomatized set theory).
- Our main goal has been to emphasize that the undecidability and incompleteness phenomenon is to be found everywhere in physics and certainly in other mathematized disciplines, and that a metamathematical exploration of those phenomena may yield a rich harvest.

We believe to have fully succeeded here.

However our work presents the following shortcomings:

- We have explicitly set aside a quantum-mechanical-like treatment of quantum physics. When we deal with quantum theories, we see them as *classical* field theories. Therefore we have excluded quantum mechanical phenomena, an area that has been essential to most philosophical discussions in our century.

This is due in part to the fact that there are still some doubts concerning the rigorous formulation within standard mathematical practice of the technique of Feynman integration, so that most of the current activity in the realm of particle physics cannot be adequately formalized within ZFC according to our approach. However recent ideas [24] [30] [32] suggest that we may also incorporate those techniques into a sound axiomatic framework.

We hope to deal with quantum mechanics in our future work.

- Our axiomatic treatment requires some sort of "superselection rule" that would discard the couples  $\langle M, G \rangle$  that do not lead to Suppes predicates for actual classical physical theories. We still do not have one simple superselection rule that would mediate from the set of all couples  $\langle M, G \rangle$  to the set of physically meaningful species of structures.
- It is not clear how our approach relates to the "states" and "observables" approach to the axiomatics of quantum mechanics. We may only suggest that tripartite species of structures, like the  $\langle S, O, L \rangle$  triple for quantum physics, and the  $\langle M, \Delta, \rho \rangle$  for an arbitrary physical theory in its relation to the real world reminds one of the three basic Bourbaki "mother"

structures, the topological, the algebraic and the order-theoretic species of structures.

### *Acknowledgments*

The questions discussed in the present paper have been elaborated by the authors since 1987 in a series of seminars presented at The University of São Paulo's Institute for Advanced Studies, and later at the Federal University of Rio de Janeiro, at the Interdisciplinary Program. The paper received its final form while F. A. D. was visiting the Institute for Mathematical Studies in the Social Sciences at Stanford University as a Senior Fulbright Scholar; we acknowledge grants from CNPq, the Fulbright Senior Scholar Program, and FAPESP.

F. A. D. wishes to thank Professor Patrick Suppes for his hospitality and many interesting comments on the subjects dealt with in this paper. The authors must also thank Professor O. T. Alas for a detailed criticism of the present paper.

Finally they wish to thank D. Getschko at the BRFPESP node of BIT-NET, who kindly helped in smoothening the communication between both authors.

### *References*

- [1] V. Arnold, *Les Méthodes Mathématiques de la Mécanique Classique*, Mir, Moscow, 1976.
- [2] M. F. Atiyah, *Geometry of Yang-Mills Fields*, Lezione Fermiane, Pisa, 1979.
- [3] J. L. Bell, *Boolean-Valued Models and Independence Proofs in Set Theory*, Clarendon Press, 1985.
- [4] N. Bourbaki, *Théorie des Ensembles*, Hermann, Paris, 1957.
- [5] N. Bourbaki, *Theory of Sets*, Hermann / Addison-Wesley, Boston, 1968.
- [6] Y. M. Cho, *J. Math. Phys.* 18 (1975), 2029.
- [7] N. C. A. da Costa / R. Chuaqui, *Erkenntnis* 29 (1988), 95.
- [8] N. C. A. da Costa / F. A. Doria, "Structures, Suppes Predicates and Boolean-Valued Models in Physics", to appear in J. Hintikka, ed. *Festschrift in Honor of V. I. Smirnov on his 60th Birthday*, 1990.
- [9] N. C. A. da Costa / F. A. Doria, "Jaśkowski's Logic and the Foundations of Physics", preprint, 1990.
- [10] N. C. A. da Costa / F. A. Doria, "Undecidability and Incompleteness in Classical Mechanics", preprint, 1990.
- [11] N. C. A. da Costa / F. A. Doria / J. A. de Barros, *Int. J. Theor. Phys.* 29 (1990), 935.
- [12] N. C. A. da Costa / F. A. Doria / J. A. de Barros, "Suppes Randomness for Arbitrary Spaces", preprint, 1990.

- [13] N. C. A. da Costa / S. French, "The Model-Theoretic Approach in the Philosophy of Science", *Philosophy of Science* 57 (1990), 248.
- [14] M. L. Dalla Chiara / G. Toraldo di Francia, *Le Teorie Fisiche*, Boringhieri, 1981.
- [15] M. Davis, *Amer. Math. Monthly* 80 (1973), 233.
- [16] M. Davis / Yu. Matijašević / J. Robinson, "Hilbert's Tenth Problem. Diophantine Equations: Positive Aspects of a Negative Solution". In: F. E. Browder, ed., *Mathematical Developments Arising from Hilbert Problems*, Proc. Symp. Pure Math., 28 (1976).
- [17] J. Dell / L. Smolin, *Commun. Math. Phys.* 65 (1979), 197.
- [18] F. A. Doria, *Lett. Nuovo Cim.* 14 (1975), 480.
- [19] F. A. Doria, *J. Math. Phys.* 18 (1977), 564.
- [20] F. A. Doria, *Commun. Math. Phys.* 79 (1981), 435.
- [21] F. A. Doria / S. M. Abrahão / A. F. Furtado do Amaral, *Progr. Theor. Phys.* 75 (1986), 1440.
- [22] F. A. Doria / A. F. Furtado do Amaral / J. A. de Barros, "Non-Curvature Solutions for the Differential Bianchi Conditions", preprint, 1990.
- [23] G. Emch, *Mathematical and Conceptual Foundations of 20th Century Physics*, North-Holland, Amsterdam (1984).
- [24] R. Henstock, *Proc. London Math. Soc.* 27 (1973), 317.
- [25] J. M. Jauch, *Foundations of Quantum Mechanics*, Addison-Wesley, 1968.
- [26] F. Klein, *Le Programme d'Erlangen*, French translation, Gauthier-Villars, Paris, 1974.
- [27] S. Kobayashi / K. Nomizu, *Foundations of Differential Geometry*, I, J. Wiley, 1963.
- [28] D. H. Krantz / R. D. Luce / P. Suppes / A. Tversky, *The Foundations of Measurement*, I, Academic Press, New York, 1971.
- [29] J. L. Krivine, *Théorie Axiomatique des Ensembles*, PUF, Paris, 1969.
- [30] P. Muldowney, *A General Theory of Integration in Function Space*, Pitman Research Notes in Math. (1987), #, 153.
- [31] L. Narens / R. D. Luce, *Psych. Bull.*, 99 (1986), 166.
- [32] L. Peng-Yee, *Lanzhou Lectures in Henstock Integration*, World Scientific, London, 1989.
- [33] D. Richardson, *J. Symbol. Logic* 33 (1968), 514.
- [34] J. Shoenfield, *Mathematical Logic*, Addison-Wesley, 1967.
- [35] W. Stegmüller, *Theorie und Erfahrung*, I, Springer, Berlin, 1970.
- [36] W. Stegmüller, *Theorie und Erfahrung*, II, Springer, Berlin, 1973.
- [37] W. Stegmüller, *The Structuralist View of Theories: A Possible Analogue of the Bourbaki Programme in Physical Sciences*, Springer, Berlin, 1979.
- [38] S. Sternberg, *Lectures on Differential Geometry*, Prentice-Hall, 1964.
- [39] P. Suppes, *Set-Theoretical Structures in Science*, mimeo., Stanford University, 1967.
- [40] P. Suppes, *Scientific Structures and their Representation*, preliminary version, Stanford University, 1988.
- [41] Private remark by P. Suppes.
- [42] C. H. Taubes, *J. Diff. Geometry* 25 (1987), 363.
- [43] A. S. Wightman, "Hilbert's Sixth Problem: Mathematical Treatment of the Axioms of Physics". In: F. Browder, ed. *Mathematical Developments Arising from Hilbert Problems*, Proc. Symp. Pure Math. 28 (1976).

# Mathematics in Philosophy

COLIN HOWSON (London)

## 1. Introduction

Mathematics has long been the object of philosophical study, both for its own sake, engendering the discipline called philosophy of mathematics, and also for the light that it throws on philosophy itself. There is no analogous discipline called the mathematics of philosophy, but that is nevertheless what I shall talk about here – about the way or ways in which mathematics has been applied to, or used in, philosophy.

Metaphysics has notoriously been prone to draw inspiration from mathematics. The temptation to see in the triad of Truth, Beauty and Number aspects of some mystical One, for example, has proved irresistible to many a metaphysician from Pythagoras and Parmenides to present-day speculative cosmologists – though the latter would perhaps prefer to replace 'Number' by 'symmetry group' (in unwitting anticipation Plato, in a late dialogue, identifies Truth, Beauty and Symmetry as the joint cause of the Good (*Philebus* 65a)). It is not the debt of metaphysics to mathematics that I am going to discuss here, however, but that of the other main branch of philosophy known as *epistemology*, or the theory of knowledge.

## 2. Rationalism and Empiricism

There are famously two great opposing theories in epistemology. These are *rationalism* and *empiricism*. Rationalism is the doctrine that all genuine knowledge is the deductive closure of a set of first principles whose necessity is grasped by a purely intellectual intuition owing nothing to the data of sensory perception. Indeed, sense perception is not only denied epistemic authority by rationalists; it is also typically cast in the negative role of adulterer and distorter of true knowledge. Empiricists, by contrast, claim that factual knowledge is authenticated by experience and only by experience. To moderns empiricism has all the plausibility. Yet philosophical rationalism was on balance the prevailing influence over the two thousand or so years from the founding of the

Academy to the end of the seventeenth century; only after that was it definitely superseded by empiricism. But mathematics was the inspiration for rationalism, and still poses the principal objection to a full-blown empiricism.

Let us deal with these issues in turn. The reason why rationalism was so compelling to its principal founding father, Plato, is to be found in the mathematics of his day. Of this he had first hand acquaintance in the proceedings of the Academy, the contemporary centre of mathematical research which he himself founded, and which numbered among its members two of the greatest of the ancient Greek mathematicians, Theaetetus and Eudoxus. Much of the Academy's research was personally directed by Plato, and the axiomatising, anti-empiricistic trend in fourth century Greek mathematics seems to have been largely due to his influence (Karasmanis [1987] p. 250: for this and other information about the relation of Plato to his contemporary mathematics I am indebted to Karasmanis' work).

The mathematical influence is apparent throughout the Dialogues, stronger in some than others; it is prominent, for example, in the *Meno*, where mathematics is set up as a model of knowledge and the Platonic doctrine of recollection exhibited by means of a geometrical problem. If we are prepared to deny ourselves the wisdom of hindsight it isn't too difficult to appreciate why mathematics should have come to be regarded as paradigmatic for knowledge. Greek mathematics, especially after it had been collated and systematised, towards the end of the fourth century, by Euclid (who was himself, according to Proclus, a disciple of the Platonic school), appeared to supply factual information with both certainty and exactness, in a way that seemed in principle not to be achievable by observation. One can be certain, to take a simple example, that the angle sum of a Euclidean plane triangle is equal to two right angles, not because of any record of measurements of particular triangles, but because it is *proved* to follow from the essential properties of a plane triangle. By contrast, no number, however large, of observations could ever render this relationship certain, for there would always remain unexamined triangles; moreover, the *exact* magnitude of the sum of the angles could never be revealed empirically either, since imperfections in the drawings would generate, at best, a scatter around the true value. So the empirical approach could achieve neither the exactness nor the necessity of this relationship revealed so elegantly *a priori*. It is hardly surprising that Plato enjoined the study of mathematics as an essential preliminary to philosophy: it taught important epistemological lessons. In particular it seemed to show that observation yields, in the Platonic terminology, mere *doxa*, opinion, while the intellect alone, unaided or rather unhindered by the senses, could render up the pure distillate of *episteme*.

*Line, plane, point, triangle* are universals. So are *man, state, justice, table* etc.. It was not at all obvious then that the incredibly fruitful methods of contemporary mathematics should not be extended to elicit with equal certainty the natures of the members of the second group – that is to say, of everything, or rather, if you are Plato, of the Idea of everything. The programme Plato inaugurated in the fourth century B.C. was to do precisely this, using in particular the method he termed *dialectic* i.e. indirect proof or *reductio*, which he regarded as preeminently the method of proof in mathematics (van der Waerden [1963] p. 149). The eventual result of the dialectical process, Plato believed, would be a systematically unified, complete *a priori* knowledge of the hierarchy of Ideas. Adumbrating such a programme in the later *Dialogues*, Plato bequeathed a methodology for knowledge which was found to be just as compelling in the seventeenth century A.D. as it was in the fourth century B.C.: to regard as the domain of truth not the world of sensible objects but that of ideas, whose structure is revealed by means of mathematical or quasi-mathematical proof.

This is, however, to anticipate. The Platonic doctrine of Ideas was, it is well-known, explicitly rejected by Plato's greatest pupil, Aristotle, who had none of Plato's contempt for observation – quite the reverse, in fact – but little of Plato's reverence for mathematics either. Aristotle taught in his own Academy, the Lyceum, but his and Plato's conflicting views were to find a sort of resolution from without, so to speak: each was made to subserve Christian dogma. Platonism was absorbed into Christian theology via Plotinus and Saint Augustine while, much later, in the thirteenth century, Aristotle was sanitised and repackaged in the work of Albertus Magnus and Aquinas. Secular learning, in Europe at any rate, revived only in the Renaissance, when Plato was rediscovered and read with appreciation, though for his style now rather than his content.

Nevertheless the Platonist programme was revived in the seventeenth century, in their different ways, by the three great rationalists of the time, Descartes, Leibniz and Spinoza. Each took mathematics, the axiomatised geometry and number theory of Euclid's *Elements*, as his model of knowledge (its influence is explicit also in Pascal's fragment *de l'esprit géométrique*, in which deductive geometry is held up as "the correct method of reasoning about all things"). Just as did Plato, Descartes thought of all science as akin to mathematics and, in the unfinished *Regulae ad Directionem Ingenii*, actually called the methodology he proposed "Universal Mathematics". Leibniz wanted his proposed Encyclopaedia cast in Euclidean form, said that his "metaphysics is all mathematics", and in true Platonic style disparaged experience as "confused thinking". Spinoza had promised in *De Intellectus Emendatione* that he would

"write about human beings as though I were concerned with lines, planes and solids", and in the *Ethics* he carried the programme through. He too denied any authority to sense experience: only knowledge of what he called the second and third kinds, which is purely intellectual, can be true according to Spinoza. This corresponds exactly to Plato's notion of *episteme*. Spinoza's knowledge of the first kind is Plato's *doxa*: it is "confused and fragmentary", indeed strictly false in Spinoza's system, and includes all of the data of sense-experience.

The geometrical model proved however, as we now know, or as most of us think we now know (though there are some notable exceptions), to be the wrong one for empirical science – as the contemporary terminology implies. To some extent this realisation grew from a deeper understanding of deductive logic. Deduction is never content-increasing, in the first place, so all that you get out of a deductive inference is what, in a sense, you put in in the beginning. And secondly, either the premisses of such an inference are purely definitional, in which case what you put in at the beginning is effectively nothing except a convention about how terms are to be understood, or they're not definitional, in which case what you put in at the beginning can never be more certain, except in a purely subjective sense, than what you infer.

Also, in the seventeenth century mathematics itself ceased to be just the contents of Euclid's *Elements*. With the piecemeal growth of analysis it was far from clear that the new mathematics could be cast as an *a priori* deductive science at all. These facts by themselves are not enough, however, to deter the convinced rationalist. The axiomatising tendency returned in the nineteenth century and remains with us, while the attempts of J. S. Mill and others in the nineteenth century to exhibit an empirical nature of mathematics completely failed, by common consent. Our knowledge of the integers *does* seem to be obtained *a priori*, and most people would probably add that it is obtained by deduction from axioms both *a priori* necessary and synthetic, being those of either Peano's Axioms or one of the current axiomatisations of set theory.

At any rate, the apparently undeniable *a priori* character of mathematical knowledge is a fact which continues to vex empiricists, whose initial strategy, adopted first by Hume and continued up to Wittgenstein and Russell, for explaining away that *a priori*, certain character was to declare mathematics void of content, mere logical truth. This is, of course, the doctrine of *logicism*. While, however, the reduction to set theory can be and was successfully achieved, set theory itself, or at any rate any of the axiomatic set theories current today, is not so easily regarded as logic. For these set theories make strong existence claims – even, including as they do an axiom of infinity, unconditional existence claims. On the other hand, if existence claims were to



be the touchstone of what is logic and what isn't, then standard (first and higher order) logic itself would deserve to be cast beyond the pale, for it asserts, and unconditionally too, that *something* exists ( $\vdash \exists x(x=x)$ ).

Nor, according to some recent analyses, is it necessary for mathematics to make existence claims at all, or at any rate claims to actual existence. One recently canvassed theory is that the natural numbers, the real numbers, and sets are not actual existents, but merely characterise *possible* structures which are unique in the sense that they are determined up to isomorphism by appropriate second order formalisations (while second order Zermelo-Fraenkel set theory is not categorical in the strict sense, nonetheless for any two full models one is isomorphic to an end extension of the other). This fact has been taken by some to support a neologicist position. For the categoricity of some second order systems means that in such cases a statement of the form "A is true in the standard model" is equivalent to the validity in second order logic of a statement of the form " $S \rightarrow A$ ", where S is the (finite) conjunction of a set of second order sentences. Thus an apparently Platonistic assertion about, say, natural numbers seems to be replaceable by a truth of logic, which Boolos for one regards as "a partial vindication" of logicism ([1975] p. 1). Hellman, building on earlier ideas of Putnam, exploits the categoricity of second order systems to take the possibilist line further. Prefixed by a necessity operator " $S \rightarrow A$ " is a truth of second order modal logic. So granted merely the *possible* existence of a model of S we seem to have a rigorous justification of the thesis that the truth of A is equivalent to the nonvacuous truth of the statement that A would hold in any structure of the relevant type. This position Hellman calls modal-structuralism (I have slightly oversimplified his account: the modal-structural translate of "A is true" is not " $\Box(S \rightarrow A)$ " but " $\Box \forall X(S \rightarrow A)^X$ ", where all the constants in A have been replaced by variables and quantified, and the relativisation to the class variable X corresponds to the explicit mention of the possible domains in which S is satisfied).

There have been other recent and much-discussed attempts to remove mathematics from its apparently natural habitat in the synthetic *a priori*. Whatever the rights and wrongs of these antiplatonist approaches (a good critical discussion of Boolos' and Hellman's stes can be found in Hanson [1990]), it remains a fact that mathematics tends anyway to be regarded as a special case by most empiricists, for whom the principal objection to rationalism has always been that all the attempts to excogitate necessary nonmathematical truths *a priori* never seemed to come up with *useful* information, that is to say information from which reliable *predictions* could be made: as Bacon never tired of pointing out, such knowledge, if knowledge it was, was sterile. The deliverances of experience, if less perfect and less certain, seemed at any rate

more fruitful. People began to note that careful measurement, admittedly guided by theory, revealed hitherto unsuspected regularities among phenomena. It became evident that some of these regularities, involving the humble pendulum as well as the planets themselves, seemed even to attain to the status of laws. Kepler's intellectual evolution epitomises the seventeenth century revolution in epistemology: starting out convinced that the five regular solids determined celestial motions, in conformity with an *a priori* Platonist epistemology, he ended by fitting curves to Tycho Brahe's observations. By the time of the publication of Newton's *Principia* that revolution was all but complete, and Newton himself used Kepler's fitted curves, the confocal ellipses, as primary data, and showed that they determined an inverse square force law. This is not to say that there are not strong *a priori* elements still in, say, Galileo and Newton. There are; but it is also clear that for them the ultimate tribunal is experience.

Despite Kant's desperate rearguard action at the end of the eighteenth century, the idea of a purely *a priori* justification for science became relegated to intellectual history. Famously, even the mathematical underpinnings of science like the Euclidean geometry which Kant regarded as constitutive of our very idea of space, became regarded as revisable, and were revised, in the light of experience. Today the real number system, classical set theory and logic have all been questioned as to their suitability as a foundation for physical theory, and non-classical theories like topos theory seriously investigated as possible replacements. The lack of immunity from what is after all empirically-based criticism that even these fundamental classical theories appear to suffer seems to me to support the view that mathematics *is* really empirical, though in the indirect way suggested by Quine in 'Two Dogmas of Empiricism' rather than through any direct correlation of its referential terms with observables. But this indirect relation with experience is nevertheless, as Quine pointed out, characteristic of physical theory in general.

### 3. Probabilism

It is time to return to the main theme of this paper. The transition from rationalism, with its role model of axiomatised geometry, to empiricism did not spell the end of the dependence of epistemology on contemporary mathematics. Far from it. Indeed, with the advent of empiricism, we see the mathematicians taking a keener interest than ever before in epistemology, an interest which they have never since lost. For the remarkable fact was that while extrapolations from experience could only ever be more or less certain, and so never attain to the status of *episteme*, the mathematicians seemed by

the end of the seventeenth century to have discovered the basic laws, mathematical in form, of *uncertainty itself*. Those basic laws were the laws of the calculus of probabilities. One of the mathematical discoveries of the seventeenth century had been combinatorial algebra, and one of its earliest applications was to the calculation of the fair odds on the various outcomes of simple games of chance like throwing dice. Probability, as we all know, is a simple mathematical transform of odds. So was born mathematical probability, and the inventor of Pascal's triangle is credited by another great probabilist with having solved the first problem of mathematical probability ('un problème relatif aux jeux de hasard, proposé à un austère janséniste par un homme du monde a été l'origine du calcul des probabilités', Poisson [1837], p. 1).

Leibniz, rationalist at bottom though he was, also took a keen interest, and so did his friend and correspondent, James Bernoulli, who proved the first great limit theorem of mathematical probability, and the fourth part of whose *Ars Conjectandi* is nothing less than a manifesto on behalf of the new science. The eighteenth century English clergyman Thomas Bayes successfully – or so it seemed – "inverted" Bernoulli's Theorem to determine the precise a posteriori probability of a simple type of statistical hypothesis, and in so doing provided an apparently definitive refutation of Hume's thesis that any argument from past to future must implicitly presuppose what it sets out to prove.

Bayes' result was further extended and given an analytic proof by Laplace, by which time it seemed that nothing but mathematical intractability prevented the calculation of the *a posteriori* probability of any hypothesis whatsoever. Mathematical empiricism had come of age. Not only had the formal laws of uncertainty been found, and transformed into a computational calculus, but so too, in the new *decision theory* based on the derivative notion of mathematical expectation, had a method of determining the relative merits of different courses of action whose inputs were the utilities of the various possible outcomes of those actions and the probabilities of the possible states of the world which generated them. Uncertainty, therefore, while remaining a disagreeable but inevitable companion of extrapolations from observations, could, it seemed, at any rate be tamed and itself made the subject of a higher-level, mathematical certainty.

Alas not. Students of the history of probability will know that this programme for a global science of uncertainty, which seemed at one point likely to fulfill Leibniz's dream of turning every question of what should be done into one simply of calculation, quickly degenerated into a prolific source of paradox and outright inconsistency, and that by the end of the first couple of decades of this century it was thoroughly discredited. One of its fundamental

principles, the Principle of Insufficient Reason, *aliter* the Principle of Indifference (Keynes' name), was the cause of the trouble. This asserted that if  $h_1, \dots, h_n$  are  $n$  exclusive and exhaustive possibilities, all 'equally possible' *a priori*, then the unconditional probability of each is  $1/n$ . In many problems an underlying metric seemed to supply a natural criterion of "equal possibility"; for example, if all one knows is that the possible values of a random variable  $X$  exhaust a bounded interval  $I$  of real numbers, then the elements of any partition of  $I$  into equal subintervals are plausibly "equally possible". A (usually tacit) continuity assumption then yields the result that the *a priori* probability that the value of  $X$  lies in any given subinterval is proportional to the length of the subinterval.

The Principle of Indifference was indispensable in generating the so-called prior distributions to be plugged into Bayes' Theorem to generate posterior distributions from observational data. It is surprising, with hindsight, that it took so long for its problematic nature to be grasped. For it generates inconsistencies with alarming ease. For example, suppose the interval  $I$  above is positive and define a new variable  $Y = X^2$ . Clearly, given that  $X$  is non-negative,  $X$  and  $Y$  are related by a continuous one-one transformation; any information which can be conveyed about the data-source by stating the observed values of  $X$  can be equivalently conveyed by stating the corresponding values of  $Y$ . It is, therefore, ultimately a matter of convention which variable is used to describe the observed phenomena. It is quite clear that  $X$  and  $Y$  cannot both be uniformly distributed in their respective intervals, but the Principle of Indifference appears to require both to be. Bertrand's famous "paradoxes of geometrical probability" arise in just this way.

The situation is essentially the same with respect to any partition of the space of possibilities, whether induced by the values of a continuously distributed variable or not. No probability distribution can be uniformly neutral over all partitions, but an *a priori* choice of any one as the set of "equal possibilities" is bound to be quite arbitrary, simply because it is *a priori*. This was the consensus at the end of the second decade of the present century, and though Carnap attempted to revive the programme in the mid fifties the majority opinion today is still that the enterprise is hopeless.

#### 4. Subjective Probability and a New Logic of Consistency

However, in the 1920's and 1930's there was a remarkable new development, which left the syntax of probability intact, but radically reinterpreted. Again, the mathematicians led the way, F. P. Ramsey in England and B. de Finetti in Italy. They independently proved a result which though mathematically simple

is of far-reaching significance. Suppose that two people A and B bet in the following way with respect to the occurrence of an event described by  $h$ . If  $h$  is true A receives from B the sum  $S(1-p)$ , where  $S$  is some positive utility, while if  $h$  is false A gives B the sum  $pS$ . So A is betting on  $h$  at odds  $p/(1-p)$ , and B against  $h$  at the reciprocal odds. Accordingly  $p$  is called the *betting quotient* on  $h$ , and  $S$  the *stake*. Ramsey and de Finetti showed that if A bets against B on any set of hypotheses, and an independent umpire determines who bets on and who against at each bet, and also the stakes, then if the betting quotients do not satisfy the finitely additive probability calculus then stakes can be set such that either A or B, named in advance, is made to suffer an inevitable loss. This is called the Dutch Book Argument. Conversely, if the betting quotients do satisfy the probability calculus, no such selection of stakes is possible.

The significance of this result for epistemology is as follows. Suppose that we define a betting quotient on  $h$  to be *fair* just in case it gives no calculable advantage to either side of a bet on  $h$  at the associated odds. If we assume, which seems plausible, that no *set* of fair odds can generate a net positive gain for either party predictable in advance, then the Dutch Book Argument shows that you cannot consistently claim that a set of betting quotients is fair which does not satisfy the probability calculus. Now if we also grant that a natural measure of your degree of belief in the truth in  $h$  is the betting quotient which you think fair for  $h$  in the light of your current knowledge, then we can infer that any consistent distribution of degrees of belief over a set of hypotheses must satisfy the probability calculus. Thus a corollary of Ramsey's and de Finetti's result is that the probability calculus furnishes the logic of consistency for partial beliefs. Nothing like the Principle of Indifference warrants inclusion among these logical principles, which alone set the standard for valid inductive inferences, characterised as transitions from a prior belief distribution to one conditional on the new observational data. The proponents of this point of view, the Personalist Bayesians, claim that with these results we can now at last, after three centuries' struggling with the problem, understand fully why probability is, as philosophers and mathematicians from James Bernoulli and Leibniz onwards have frequently claimed it is, the *logic* of induction.

## 5. Conclusion

The Personalist Bayesian theory is not the only theory of inductive inference, though it continues to enlist more support than its more recent competitors (a recent detailed account is Howson and Urbach [1989]). Of these the most

popular are Dempster-Shafer theory, in which belief functions are defined as (a subclass of) lower probabilities, and hence are nonadditive (Shafer [1976]), and the fuzzy probability and logic approach (a useful recent survey is provided in Dubois and Prade [1989]). Then there are other theories which are not even nominally probabilistic, for example that recently proposed by Spohn based on so-called natural conditional functions (Spohn, forthcoming). Different though these theories of uncertain inference are from each other, they share a common feature which has come so much to be taken for granted that it tends to pass without comment: they are all overtly *mathematical* theories. Just as it is now inconceivable that physics can be done except in the context of some mathematical theory or other, so too is that becoming true of the empiricist epistemology to which, if not all, the vast majority now subscribe.

### References.

- Boolos, G. (1975), "On Second Order Logic". In: *Journal of Philosophy*, 72, 509–523.
- Dubois, D. / Prade, H. (1989), "Modelling Uncertainty and Inductive Inference: A Survey of Recent Non-additive Systems". In: *Acta Psychologica* 68, 53–78.
- Hanson, W. H. (1990), "Second Order Logic and Logicism". In: *Mind* 99, 90–99.
- Hellman, G. (1989), *Mathematics without Numbers*, Oxford: The Clarendon Press.
- Howson, C. / Urbach, P. M. (1989), *Scientific Reasoning: the Bayesian Approach*, La Salle: Open Court.
- Karasmanis, V. (1987), *The Hypothetical Method in Plato's Middle Dialogues*, unpublished D. Phil. dissertation, Brasenose College, Oxford.
- Poisson, S. D. (1837), *Recherches sur la probabilité de jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilités*, Paris.
- Shafer, G. (1976), *A Mathematical Theory of Evidence*, Princeton: Princeton UP.
- Spohn, W. (forthcoming): "A General Non-probabilistic Theory of Inductive Reasoning". In: *Proceedings of the 1990 Workshop in the Foundations of Probability*, Paris.
- van der Waerden, B. L. (1963), *Science Awakening: Egyptian, Babylonian and Greek Mathematics*, New York: John Wiley.



## Historical Dimensions





## Are There Revolutions in Mathematics?

JOSEPH W. DAUBEN (New York)

Revolutions never occur in mathematics.

Michael J. Crowe

The calculus – it all embellishes the spirit and has created, in the world of geometry, an unmistakable revolution.

Bernard de Fontenelle

Nonstandard analysis is revolutionary. Revolutions are seldom welcomed by the established party, although revolutionaries often are.

G. R. Blackley

In the 1870's the German mathematician and historian of mathematics Hermann Hankel characterized what he took to be the essence of mathematics in very structural terms. In trying to express how mathematics develops, he used the following concrete metaphor:

In most sciences one generation tears down what another has built, and what one has established, another undoes. In mathematics alone each generation builds a new story to the old structure.<sup>1</sup>

Based upon such views, it has often been argued that revolutions do not occur in mathematics, and that unlike the other sciences, mathematics accumulates *positive knowledge* without revolutionizing or rejecting its past. The "old structures" are simply embedded in later additions; nothing is ever lost in the

---

<sup>1</sup> H. Hankel, *Die Entwicklung der Mathematik in den letzten Jahrhunderten, Antrittsvorlesung* (Tübingen, 1871; 2nd ed. 1889), p. 25. Similar views have also been voiced by G. D. Birkhoff, "Mathematics: Quantity and Order", *Science Today* (1934), pp. 293–317, esp. p. 302, and *Collected Mathematical Papers*, New York: American Mathematical Society, 1950, vol. III, p. 557; and C. Truesdell: "While 'Imagination, fancy, and invention' are the soul of mathematical research, in mathematics there has never yet been a revolution". In: *Essays in the History of Mechanics*, New York: Springer-Verlag, 1968, foreword.

house of cards that mathematicians are forever extending, but, apparently, never rebuilding.

Once established, new theorems and results become a part of a mathematical continuum – forever.<sup>2</sup> Unlike the physical and life sciences, littered with the names of historical relics whose theories were found inadequate or even completely discarded in later times – mathematics is different. Names like Pythagoras, Eudoxos, Euclid, Archimedes, Apollonios, Diophantos – may have a dusty ring to them, but their results seem timeless and their works may still be read with profit and not solely for their historical interest.

Emphatic views on the nature of mathematics and the problem of revolutions were expressed not long ago by the American historian of mathematics, Michael Crowe (speaking at a Workshop on the Evolution of Modern Mathematics held at the American Academy of Arts and Sciences in Boston, August 7–9, 1974). In a short paper on "Ten 'Laws' Concerning Patterns of Change in the History of Mathematics", (subsequently published in *Historia Mathematica*), he concluded emphatically with his "tenth law" that "Revolutions never occur in mathematics".<sup>3</sup>

### 1. *Revolutions and the History of Mathematics*

Whether revolutions can be discerned in any discipline depends, of course, upon one's definition of "revolution". In insisting that "revolutions never occur in mathematics", Crowe explains that his reason for asserting this tenth

---

2 Claude Bernard regarded mathematics as essentially different from the sciences for exactly this reason: "mathematical truths are immutable and absolute", he insisted; "mathematics grows by simple successive juxtaposition of all acquired truths". In contrast, the experimental sciences are only relative, consequently they "can move forward only by revolutions and by recasting old truths in a new scientific form". See C. Bernard, *An Introduction to the Study of Experimental Medicine*, trans. H. C. Greene, New York: The Macmillan Co., 1927; reprinted New York: Dover, 1957, p. 41. Bernard's views on the subject of scientific revolutions are described in Supplement 5.3, "Revolution in Mathematics", in: I. B. Cohen, *Revolution in Science*, Cambridge, Mass.: The Belknap Press of Harvard UP, 1985, pp. 488–491, esp. p. 491.

3 M. J. Crowe, "Ten 'Laws' Concerning Patterns of Change in the History of Mathematics", *Historia Mathematica*, 2 (1975), pp. 161–166, esp. p. 165; an even shorter version was published as a contribution to the "Workshop on the Evolution of Modern Mathematics", held at the American Academy of Arts and Sciences in Boston, August 7–9, 1974. See M. Crowe, "Ten 'Laws' Concerning Conceptual Change in Mathematics", *Historia Mathematica*, 2 (1975), pp. 469–470, esp. p. 470.

"law" depends upon his own definition of revolutions. As he puts it: "My denial of their existence (revolutions) is based on a somewhat restricted definition of 'revolution', which in my view entails the specification that a previously accepted entity *within* mathematics proper be rejected".<sup>4</sup> (Yet, having said this, Crowe is nevertheless willing to admit that non-Euclidean geometry, for example, "did lead to a revolutionary change in views as to the nature of mathematics, i.e. a revolution in the philosophy of mathematics, but not *within* mathematics itself").<sup>5</sup>

Certainly one can question the definition that Crowe adopts for "revolution". It is unnecessarily restrictive, and in the case of mathematics, it automatically eliminates "revolutions" altogether. Because of the narrow conception of the term, revolutions become inherently impossible within his conceptual framework. But before deciding whether there are grounds for challenging Crowe's "Tenth Law", it may be helpful to consider briefly the meaning of "revolution" as an historical concept.

In fact, the term first made its appearance with reference to scientific and political events in the 18th century, although with considerable confusion and ambiguity as to the meaning of the term in such contexts. In general, the word "revolution" was regarded in the 18th century as indicating a decisive breach of continuity, a change of great magnitude – even though the old astronomical sense of revolution as a cyclical phenomenon persisted as well. But following the French Revolution, the new meaning gained currency, and thereafter, revolution commonly came to imply a radical change or departure from traditional or acceptable modes of thought.<sup>6</sup> Revolutions, then, may be visualized as a series of discontinuities of such magnitude as to constitute definite breaks with the past. After such episodes, one might say that there is no returning to an older order.

Bernard de Fontenelle, nephew of the playwright Pierre Corneille and *secrétaire perpétuelle* of the French Académie des Sciences, may well have been the

---

4 Crowe 1974, p. 470. Recently, Caroline Dunmore has written along similar lines, arguing that revolutions may occur in metamathematics, but not in mathematics proper. Refer to C. Dunmore, "Revolutions at the Meta-Level: Negative Numbers, Complex Numbers and Quaternions". In: Donald Gillies, ed., *Revolutions in Mathematics*, Oxford: Oxford UP, in press.

5 Crowe 1974, p. 470.

6 I. B. Cohen, "The Eighteenth-Century Origins of the Concept of Scientific Revolution", *Journal of the History of Ideas*, 37 (1976), pp. 257–288. Consult as well *The Newtonian Revolution, with Illustrations of the Transformation of Scientific Ideas*, Cambridge: Cambridge UP, 1980, esp. Chapter 2, pp. 39–49; and Chapter 4, "Transformations in the Concept of Revolution". In: Cohen 1985 [note 2 above], pp. 51–76.

first author to apply the word "revolution" to the history of mathematics, and specifically to its dramatic development in the 17th century. What he perceived, thanks to the infinitesimal calculus of Newton and Leibniz, was a change of so great an order as to have altered completely the state of mathematics.<sup>7</sup> In fact, Fontenelle went so far as to pinpoint the date at which this revolution had gathered such force that its effect was unmistakable. In his *éloge* of the mathematician Michel Rolle, Fontenelle referred to the work of the Marquis de l'Hôpital, his *Analyse des infiniment petits* (first published in 1696, with later editions in 1715, 1720, 1768), as follows:

In those days the book of the Marquis de l'Hôpital had appeared, and almost all the mathematicians began to turn to the side of the new geometry of the infinite, until then hardly known at all. The surpassing universality of its methods, the elegant brevity of its demonstrations, the finesse and directness of the most difficult solutions, its singular and unprecedented novelty, it all embellishes the spirit and has created, in the world of geometry, an unmistakable revolution.<sup>8</sup>

In describing the "new geometry" as bringing about an "unmistakable revolution", it was also a "quantum leap" from the old mathematics in the sense that a theory of infinitesimals and limits was not part of previously accepted branches of mathematics – in particular Euclidean geometry, number theory, or even analytic geometry. The "new geometry" it might be said, constituted a breach requiring a leap of faith to new comprehensive concepts and methods that were not part of earlier mathematical practice.

Moreover, these were not simply accumulations or "innovations" – as opposed to the stuff of "revolutions". New additions to mathematics are made all the time, but seldom do these have so substantial an effect on the content, methods or meaning of mathematics as to constitute true revolutions. The introduction, for example, of infinitesimals brought something new that was not part of the previously accepted branches of mathematics – Euclidean geometry, number theory, or even analytic geometry. Infinitesimals represented a significant departure, as yet philosophically unjustified but surprisingly powerful in the hands of Newton, Leibniz and their followers. The new concepts were general and comprehensive, and not part of earlier mathematical practice. Nor could that practice have produced these – new elements and methods had to be

---

<sup>7</sup> Bernard de Fontenelle, *Eléments de la géométrie de l'infini*, Paris, 1727, especially the preface, which is also reprinted in Fontenelle, *Oeuvres de Fontenelle*, Paris, 1792, vol. VI, p. 43.

<sup>8</sup> B. de Fontenelle, "Eloge de M. Rolle", *Histoire de l'Académie Royale des Sciences*, Paris, 1719, pp. 84–100, esp. p. 98; see also *Oeuvres*, vol. VII, p. 67.

added to the repertoire of then-established mathematics before the revolution could occur. It was the qualitatively different mathematics introduced by the "infinitesimal calculus" that created the revolution, and no amount of ingenuity working within the confines of Euclidean geometry, number theory or analytic geometry could have brought it about.

This is borne out, in fact, in the opposition to the new methods the calculus provoked (in either its Newtonian or Leibnizian version), but more of this in a moment. Suffice it to say that the appearance of the calculus was "revolutionary", requiring the introduction of wholly new elements – having so substantial an effect on the content, methods and meaning of mathematics as to constitute a true, or as Fontenelle says, "unmistakable" revolution.

Clearly this revolution was qualitative, as all revolutions must be. It was a revolution which Fontenelle perceived in terms of character and magnitude, without invoking any displacement principle – any rejection of earlier mathematics – before the revolutionary nature of the new geometry of the infinite could be proclaimed. For Fontenelle, Euclid's geometry had been surpassed in a radical way by the new geometry in the form of the calculus, and this was undeniably revolutionary.

Traditionally, then, revolutions have been those episodes of history in which the authority of an older, accepted system has been undermined and a new, better authority appears in its stead. Such revolutions represent breaches in continuity, and are of such degree, as Fontenelle says, that they are unmistakable even to the casual observer. Fontenelle has aided us, in fact, by emphasizing the discovery of the calculus as one such event – and he even takes the work of l'Hôpital as the identifying marker, much as Newton's *Principia* of 1687 marked the scientific revolution in physics or the Glorious Revolution of the following year marked England's political revolution from the Stuart monarchy. The monarchy, we know, persisted, but under very different terms.

In much the same sense, it seems clear that revolutions *have* occurred in mathematics. However, because of the special nature of mathematics, and the unique character of its principles and the logical structure of its arguments, it is not always the case that an older order is necessarily refuted or rejected outright. Although it may persist, the old order nevertheless does so under different terms, often in radically altered or expanded contexts. Moreover, it is generally true that the new ideas would never have been permitted within the strictly construed interpretation of the old mathematics, even if the new mathematics finds it possible to accommodate the old discoveries in a compatible or consistent fashion. In most cases, many of the theorems and discoveries of the older mathematics are consigned to a significantly lesser position as a result of a conceptual revolution that brings an entirely new

theory or mathematical discipline to the foreground. This was certainly how Fontenelle regarded the calculus.

Previously I have argued this case for revolutions in mathematics with two examples:

1. The discovery of incommensurable magnitudes in antiquity, and the problem of irrational numbers that it engendered.
2. The creation of transfinite set theory and the revolution brought about by Georg Cantor's new mathematics of the infinite in the 19th century.<sup>9</sup>

Rather than reiterate the details of these two case studies again, suffice it to say that both are basic examples of significant transformations, indeed revolutions, in mathematics. In what follows, I shall instead consider three additional closely related case histories – and explain how each represents yet another example of revolutionary change in mathematics, namely:

1. The simultaneous, yet independent discovery of the infinitesimal calculus by Newton and Leibniz in the 17th century.
2. The introduction of new standards of rigor for the calculus by Augustin-Louis Cauchy in the 19th century.
3. The creation, in this century, of nonstandard analysis by Abraham Robinson.

Each of these may be regarded as more than just a new departure for mathematics. It remains to be shown how each represents a new way of doing mathematics, by means of which the face and framework of mathematics were dramatically altered in ways that indeed proved to be revolutionary.

## *2. The Calculus as Revolutionary Mathematics*

The 17th century in Europe represents a period of dramatic political, social and intellectual ferment. Mathematics as well experienced unprecedented activity, new areas of study were explored, new problems were pursued with innovative methods going well beyond the familiar, traditional bounds of axiomatic Euclidean geometry. Although the "rigor" of the ancients was a goal to which many still aspired – whenever convenient or as best they could – the most creative 17th-century minds were willing to experiment with new ideas

---

<sup>9</sup> J. W. Dauben, "Conceptual Revolutions and the History of Mathematics. Two Studies in the Growth of Knowledge", E. Mendelsohn, ed., *Transformation and Tradition in the Sciences*, Cambridge: Cambridge UP, 1984, pp. 81–103.

suitable for analyzing such problems as acceleration, projectile motion, planetary orbits, or centers of gravity, to name but a few.

In fact, the "revolutionary" character of 17th-century mathematics might be sought in a number of different areas: in such methods as the new analytic geometry pioneered by Viète and Descartes; the method of infinitesimals advanced by Cavalieri and Roberval; methods of maxima/minima; infinite series; rules for finding tangents to curves; procedures for determining lengths and areas under curves; and new approaches to determining centers of gravity. These were, for the most part, problems that had not been seriously considered in earlier periods, and the methods, often applied in ingenious ways to resolve them, were not only new, but often impressive for their universality. A premium was placed on finding solutions that were general enough to apply to the widest possible classes of problems, and of all these new methods, the calculus proved to be the most revolutionary (precisely as Fontenelle said, because of its "universality ... brevity ... and unprecedented novelty").

The "revolutionary" character of the calculus, however, does not lie simply in the method of tangents and discovery of its inverse connection with the problem of quadratures. There are many examples of mathematicians in the 17th century, including Descartes, Fermat and Barrow, who recognized the importance of discovering and applying such methods. What was needed, above all, was a truly *general* method.<sup>10</sup>

Some have suggested that Isaac Barrow was actually the first to provide such a general method and that he did so, specifically, thanks to the reciprocal connection he established in his *Geometrical Lectures* between the problem of finding the tangent to a curve and the area under it.<sup>11</sup> But as Michael Mahoney has recently argued – in terms that point to the genius of what later Newton

---

<sup>10</sup> The origins and development of the calculus have been the subject of historical study and polemic since the 17th century. Among the best (and most recent) of these are C. B. Boyer, *History of the Calculus and Its Conceptual Development*, New York: Dover, 1959; M. Baron, *The Origins of the Infinitesimal Calculus*, Oxford: Pergamon Press, 1969; C. H. Edwards, *The Historical Development of the Calculus*, New York: Springer Verlag, 1979; K. M. Andersen, "Techniques of the Calculus, 1630–1660", and H. J. M. Bos, "Newton, Leibniz and the Leibnizian Tradition", both in I. Grattan-Guinness, ed., *From the Calculus to Set Theory, 1630–1910. An Introductory History*, London: Duckworth, 1980, pp. 10–48 and pp. 49–93, respectively; and N. Guicciardini, *The Development of Newtonian Calculus in Britain, 1700–1800*, Cambridge: Cambridge UP, 1989.

<sup>11</sup> J. M. Child, *The Geometrical Lectures of Isaac Barrow*, Chicago: The Open Court Publishing Co., 1916, and "The 'Lectiones Geometricae' of Isaac Barrow", *The Monist*, 26 (1916), pp. 15–20. Consult as well Baron 1969.



and Leibniz actually *did* see – "Only to the retrospective eye", might one be misled to assume that Barrow had formulated the basic ideas of what was to become the Newton-Leibniz calculus. What Barrow failed to appreciate, however, was especially significant:

Barrow posited no general framework for generating curves by concurrent motions... he showed no hint of allowing both of the component motions (horizontal and vertical) to vary over time and to relate the resulting distances through an equation... Nor *a fortiori* did he translate the proposition into a method of tangents by reducing it to rules by which, given such an equation, one calculates the ratio of the component velocities at any point.<sup>12</sup>

It was in Proposition 16, Lecture IV, that Barrow connected tangents with quadratures. And as Mahoney says, this did "forge a link that escaped many others", doubtless because most mathematicians regarded tangents and quadratures as two separate, seemingly unrelated problems. But as Mahoney also cautions, what Barrow had formed was really "an *ad hoc* relation, tied immediately to the geometrical configuration rather than to an algebraic framework (as presented at the end of Lecture X, for example)".<sup>13</sup>

Mahoney warns that arguments favoring Barrow's priority in recognizing the significance of the calculus are the products of faulty hindsight by which Barrow's geometric propositions only seem like the calculus.<sup>14</sup> In fact, Barrow himself never pretended to have discovered a method that was in any way a central, organizing concept. Although many mathematicians have taken Barrow's Proposition 19 in Lecture X as equivalent to the Fundamental Theorem of the Calculus, as Mahoney has said, it was "clearly not fundamental for Barrow".<sup>15</sup> Above all, Barrow failed to make the fundamental connection between summations and the determination of tangents to curves as *inverse operations*. (Mahoney also notes that Barrow did not consider ratios as quantities – a step both Newton and Leibniz took with their algebraic analysis – the former introducing fluxions, the latter differentials).

Newton and Leibniz, however, did make this significant connection, and placed the calculus (that each had invented) at the center of mathematical structures of surpassing generality and power. In the course of their famous contro-

<sup>12</sup> When Barrow *did* get to such matters in Lecture X of the *Geometrical Lectures*, "it bore no relation ... to kinematic generation of curves". See M.S. Mahoney, "Barrow's Mathematics: Between Ancients and Moderns". In: M. Feingold, ed., *Before Newton: The Life and Times of Isaac Barrow*, Cambridge: Cambridge UP, 1990, pp. 179–249, esp. p. 211.

<sup>13</sup> Mahoney 1990, p. 212.

<sup>14</sup> Mahoney 1990, p. 236.

<sup>15</sup> Mahoney 1990, p. 236.

versy over who deserved priority for making this discovery, the debate itself, especially the vehemence with which it was argued on both sides of the English Channel, indicates the "revolutionary" character of the discovery. Certainly it suggests that both Leibniz and Newton (along with their various partisan supporters) realized that the calculus represented a seminal part of mathematics, a powerful new tool that both wished to claim.

None of the many (and interesting, sometimes significant) psychological, theological, philosophical or political issues that underlay aspects of the Leibniz-Newton controversy will be considered here. Instead, emphasis is limited to several fundamental questions of technical, mathematical significance about the calculus and how it may be taken to represent a "revolution" in mathematics.

As already noted, seventeenth-century mathematicians faced a wide variety of problems, many of them new, including problems of finding tangents to curves, determining lengths and areas, finding maxima and minima. The newly-discovered analytic geometry of Descartes and Fermat provided a fertile and suggestive context within which to examine and identify similar groups of curves, and to handle them mathematically with very general methods. Although Descartes, Fermat, Roberval, Hudde, and Sluse (among others) had begun the search for more comprehensive methods, in virtually every case there were problems. Either the methods were difficult or they were lacking in sufficient generality. But the time was ripe for new algorithmic insights that would unify – and simplify – work on all of these problems.

Drawing heavily on results of Sluse, Newton offered a number of "Universal Theorems" in an early manuscript on "The General Problem of Tangents and Curvature Resolved for Algebraic Curves". Among these, for example, he gave rules for finding tangents to algebraic curves through computation of the subnormal. At the same time, Newton showed how power series expansions could be used to widen the class of functions amenable to his new methods. Finally, he discovered the secret of reducing all of the major problems that had preoccupied 17th-century mathematicians to just two methods. As he put it, "All the difficulties hereof may be reduced to these two problems only, namely the inverse methods of fluxions and fluents".

Similarly, in a draft of the preface he wrote for the *Commercium Epistolicum* (Newton's anonymously published report, issued by the Royal Society in favor of his priority over Leibniz in discovering the Calculus), Newton explained that, in addition to handling quadratures, the method he had devised in 1671:

... was very general & easy without sticking at surds or mechanical curves & extended to finding tangents, areas, lengths, centers of gravity & curvatures of Curves, &c... [It also] extended to inverse Problems of tangents & others

more difficult, & was so general as to reach almost all Problems except numerical ones like those of Diophantus.<sup>16</sup>

These claims recalled others Newton had already published in his famous Scholium to the "fluxions lemma" at the beginning of Book II of his *Mathematical Principles of Natural Philosophy* (1687). There, in the Scholium to Lemma II following Proposition VII, Theorem V, Newton stressed his "method of determining maxima and minima, of drawing tangents...which served for irrational terms as well as rational ones".<sup>17</sup>

Leibniz too heard of the comprehensive scope of Newton's results in letters he received from John Collins via Henry Oldenburg as early as 1673:

Mr. Newton hath invented (before Mercator publish't his logarithmo-technica) a general method of the same kind for the quadrature of all curvilinear figures, the straightening of curves, the finding of the centers of gravity and solidity of all round solids...with infinite series for the roots of affected equations, easily composed out of those for pure powers.<sup>18</sup>

It was Leibniz's calculus, however, that was not only the first to appear, but the first to be applauded in print. In 1685 John Craige, the Scottish mathematician, having read Leibniz's papers in the *Acta Eruditorum*, wrote as follows in his *Methodus Figurarum Lineis Rectis & Curvis Comprehensarum Quadraturas Determinandi*. He had nothing but praise for Leibniz, who:

... shows a neat way of finding tangents even though irrational terms are as deeply involved as possible in the equation expressing the nature of the curve, without removing the irrationals.<sup>19</sup>

This was basically the message Leibniz himself sent to Huygens in a letter of July, 1690, in which he tried to explain the principles of his own calculus. Leibniz emphasized the way it enabled him:

16 "Newton's References to the 1671 Tract in his *Commercium*: An Extract from a Draft Preface", [Add. 3968.39:583r], reproduced in D.T. Whiteside, ed., *The Mathematical Papers of Isaac Newton, 1670-1673*, vol. III, Cambridge: Cambridge University Press, 1969, p. 20.

17 See F. Cajori, ed., *Sir Isaac Newton's Mathematical Principles of Natural Philosophy and His System of the World*, Berkeley: University of California Press, 1934; rep. 1966; pp. 253-54.

18 Letters 2196 and 2208. In: A. R. Hall and M. B. Hall, eds., *The Correspondence of Henry Oldenburg*, vol. IX, Madison: University of Wisconsin Press, 1973. Also quoted in A. R. Hall, *Philosophers at War. The Quarrel between Newton and Leibniz*, Cambridge: Cambridge UP, 1980, p. 50.

19 J. Craige, *Methodus Figurarum Lineis Rectis & Curvis Comprehensarum Quadraturas Determinandi*, London: M. Pitt, 1685, p. 27. Quoted in Hall 1980, p. 78.

...to subject to analysis that which M. Descartes himself had excepted from it ... and just as roots are the reciprocals of differences ... By means of this calculus I presume to draw tangents and to solve problems of maxima and minima, even when the equations are much complicated with roots and fractions ... and by the same method I make the curves that M. Descartes called mechanical submit to analysis.<sup>20</sup>

Leibniz, in response to the first open challenge to his priority in discovering the calculus made by Fatio de Duillier, published a "Reply" in the *Acta Eruditorum* in May, 1700. In stressing the general power of the new methods, he offered the brachistochrone as proof that the new calculus dealt with the most difficult problems with "great simplicity and generality".<sup>21</sup> Leibniz realized that his method clearly represented a wholly new system of mathematics, and it was as if he believed he had found a key, literally, opening "a door to a completely new realm of mathematical invention".<sup>22</sup>

Indeed, both Newton and Leibniz had succeeded in developing the calculus not only as a method, but a method that was both general and comprehensive in the variety of problems to which it applied. The many challenge problems issued by both sides during the Newton-Leibniz controversy, in attempts not-so-subtle to outshine each other, attest to the difficulty of individual questions, and to the power with which the calculus was often able to meet them. This was enough, namely the universality and novelty of the calculus, to convince Fontenelle that it was "revolutionary". But there are other indicators as well confirming the revolution brought about by the calculus, whether in the hands of Newton, Leibniz, the Bernoullis, or any number of their successors in the 18th century.

Thomas Kuhn in his *The Structure of Scientific Revolutions* mentions a number of "indicators" of Scientific Revolutions, including resistance to new, revolutionary theories on the one hand, and the eventual appearance of textbooks on the other which serve to endorse and disseminate the revolution.<sup>23</sup> As for the calculus, textbooks began to appear almost immediately, introducing the new methods both in England and on the continent. Among the most successful were those of Craige, the Marquise de l'Hôpital, Collins, Raphson, and Buffon, among others.<sup>24</sup>

<sup>20</sup> Leibniz in a letter to Huygens July, 1690, quoted from Hall 1980, p. 87.

<sup>21</sup> Hall 1980, p. 127.

<sup>22</sup> Hall 1980, p. 127.

<sup>23</sup> T. S. Kuhn, *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press, 1962; 2nd ed. 1970 (with a "Postscript-1969", 174-210).

<sup>24</sup> See the detailed discussion of this question in N. Guicciardini, *The Development of Newtonian Calculus in Britain 1700-1899*, Cambridge: Cambridge UP, 1989.

And of course, there was opposition too, some of it well known, as were the penetrating critiques of Bishop Berkeley, Bernhardt Nieuwentijdt and Michel Rolle.<sup>25</sup> But rather than examine either of these here – opposition to the calculus or the early textbooks – another Kuhnian indicator is perhaps of greater interest, namely the "paradigm shift" that occurred in mathematics as the calculus came to be increasingly accepted and used. In this case the evidence that the calculus was creating a revolution in mathematics is reflected in the new terminology and unfamiliar notations that were introduced for both the fluxional and the differential calculus.

As is often the case with the great revolutions in science, the revolution brought about by the calculus was both conceptual and visual. Consider, for example, the freshly minted terms and symbols, sums, differences, integrals, differentials, fluents, fluxions, pricked or dotted letters,  $dy/dx$ , and so on – all were concrete signs that the revolution had brought about a new order in mathematics. Superficially on the symbolic level, much deeper on conceptual, methodological levels, the new mathematics *looked* different, and *worked* differently as well. The language was new because the elements, problems, methods and results were all dramatically new. In either case, Newton's fluxional calculus or Leibniz's differential calculus, the revolution was rooted in new language. Both Newton's fluxions and Leibniz's differentials empowered methods with diverse applications without parallel in their generality among any of their predecessors.

This is all captured, for example, in Leibniz's algebraic generalization of the inverse tangent problem in terms embracing the tangent, subtangent and subnormal together, whereas Barrow had treated each of them separately. Leibniz's amalgamation represents, in fact, the powerful connections he succeeded in forging – but which had eluded Barrow.

Michael Mahoney summarizes the significance of this difference nicely in terms of structures:

Couching problems from different domains in the same symbolic language revealed common underlying structures and the relations of various structures to one another.<sup>26</sup>

As he goes on to say, "Barrow's propositions contained the substance, but not the concepts, of the calculus".<sup>27</sup>

---

<sup>25</sup> I. Grattan-Guinness, "Berkeley's Criticism of the Calculus as a Study in the Theory of Limits", *Janus*, 56 (1969), pp. 213–227; and Boyer 1959, p. 241.

<sup>26</sup> Mahoney 1990, p. 239.

<sup>27</sup> Mahoney 1979, p. 239–240.

The concepts of the calculus, as introduced by Newton, Leibniz and their adherents, were new. They required a special language suitably tailored for the applications they were meant to address, all of which (as Fontenelle said), constituted a true revolution in mathematics, easily recognizable by the beginning of the 18th century.

### 3. Cauchy's Revolution in Rigor

The revolution brought about by Newton and Leibniz was not without its problems, as the opposition just mentioned of Berkeley, Nieuwentijdt, Rolle, and others attests. In fact, the 18th century, despite its willingness to use the calculus, seems to have been plagued with a concomitant sense of doubt as to whether its use was really legitimate or not. It worked, and lacking alternatives, mathematicians persisted in applying it in diverse situations. Nevertheless, the foundational validity of the calculus was often the subject of discussion, debate, and prize problems. The best-known of these was the competition announced in 1784 by the Berlin Academy of Sciences. Joseph-Louis Lagrange had suggested the question of the foundations of the calculus, and the contest in turn resulted in two books on the subject, Simon L'Huilier's *Exposition élémentaire* and Lazare Carnot's *Réflexions sur la métaphysique du calcul infinitésimal*.<sup>28</sup>

Most histories of mathematics credit Augustin-Louis Cauchy with providing the first "reasonably successful rigorous formulation" of the calculus.<sup>29</sup> This not only included a precise definition of limits, but aspects (if not all) of the modern theories of convergence, continuity, derivatives and integrals. As Judith Grabiner has said in her detailed study of Cauchy, what he accomplished was an "apparent break with the past". The break was also revolutionary, not

---

<sup>28</sup> Lagrange also responded to the foundations problem, but did not submit a contribution of his own for the contest set by the Berlin Academy. Nevertheless, his own book, *Fonctions analytiques*, was designed to show how the calculus could be set on rigorous footing. Although L'Huilier won the Academy's prize, the committee assigned to review the submissions complained that it had "received no complete answer". None of the contributions came up to the levels of "clarity, simplicity and especially rigor" which the committee expected, nor did any succeed in explaining how "so many true theorems have been deduced from a contradictory supposition". On the contrary, the committee was disappointed that none of the prize papers had shown why infinitesimals were acceptable at all. For details, see J. V. Grabiner, *The Origins of Cauchy's Rigorous Calculus*, Cambridge, Mass.: MIT Press, 1981, pp. 40–43.

<sup>29</sup> Grabiner 1981, p. viii.

for what Cauchy introduced conceptually, but methodologically. As she concludes, Cauchy "was responsible for the first great revolution in mathematical rigor since the time of the ancient Greeks".<sup>30</sup>

This, presumably, is a revolution in mathematics that Crowe, for example, would accept, for Cauchy's revolution was concerned with rigor on a meta-mathematical level affecting the foundations of mathematics. But as shall be argued here, changes in foundations cannot help but affect the structures they support, and in the case of Cauchy's new requirements for rigorous mathematical arguments in analysis, the infinitesimal calculus underwent a revolution in style that was soon to revolutionize its content as well.

In order to appreciate the sense in which Cauchy's work may be seen as revolutionary, it will help to remember that for most of the 18th century (with some notable exceptions) mathematicians like the Bernoullis, L'Hôpital, Taylor, Euler, Lagrange, and Laplace were interested primarily in results. The methods of the calculus were powerful and usually worked with remarkable success, although it should be added that these mathematicians were not oblivious to questions about *why* the calculus worked or whether there were acceptable foundations upon which to introduce its indispensable, but also most questionable element – infinitesimals. Such concerns, however, remained for the most part secondary issues.

In the 19th century foundational questions became increasingly of interest and importance, in part for a reason that concerns the sociology of mathematics involving both matters of institutionalization and professionalization. As many mathematicians were faced with *teaching* the calculus, questions about how to define and justify limits, derivatives, and infinite sums, for example, became unavoidable.

Cauchy was not alone, however, in his concern for treating mathematics with greater conceptual rigor – at least when he was teaching at the Ecole Polytechnique or writing textbooks like his *Cours d'analyse de l'école polytechnique*.<sup>31</sup> Others, like Gauss and Bolzano, were also concerned with such problems as treating convergence more carefully, especially without reference to geometric or physical intuitions.<sup>32</sup> Whether or not Cauchy based his own

<sup>30</sup> Grabiner 1981, p. 166.

<sup>31</sup> This was only the first of a series of books that Cauchy produced as a result of his lectures at the *Ecole*. Among others, mention should be made of his *Résumé des leçons données à l'école polytechnique sur le calcul infinitésimal*, Paris: Imprimerie Royale, 1823; *Leçons sur les applications du calcul infinitésimal à la géométrie*, Paris, 1826–1828; and *Leçons sur le calcul différentiel*, Paris: de Bure Frères, 1829.

<sup>32</sup> What sets them apart, in fact, is that neither Gauss nor Bolzano was con-

rigorization of analysis upon his reading of Bolzano (as Ivor Grattan-Guinness has suggested), or by modifying Lagrange's use of inequalities and the development, in particular, of an algebra of inequalities (as Grabiner argues), it remains true in any case that Cauchy was the first to write textbooks that became models for disseminating the new "rigorous" calculus – and that others soon began to work in the innovative spirit of Cauchy's arithmetic rigor.<sup>33</sup>

Niels Henrik Abel was among the first to apply Cauchy's techniques in connection with his own important results on convergence. Bernhard Riemann revised Cauchy's theory of integration, and Karl Weierstrass further systematized Cauchy's work by carefully defining the real numbers and emphasizing the crucial distinctions between convergence, uniform convergence, continuity, and uniform continuity.

Much of what Cauchy accomplished, however, had been anticipated by Lagrange, perhaps much as Barrow and others had prepared the way for Newton and Leibniz. For example, Lagrange had already given a rigorous definition of the derivative – and surprisingly, perhaps, he used the now-familiar method of deltas and epsilons. Actually, the deltas and epsilons are Cauchy's, but the idea is due to Lagrange: the only symbolic difference is the fact that Lagrange used *D* (*donnée*) for Cauchy's epsilon and *i* (*indeterminée*) for Cauchy's delta. Both Lagrange and Ampère in fact used the method of inequalities as a useful method of proof, but Cauchy saw that it could also be used more essentially in definitions. As Grabiner has said, Cauchy extended this method to defining limits and continuity, and in doing so:

...achieved exactly what Lagrange had said should be done in the subtitle of the 1797 edition of his *Fonctions analytiques*; namely the establishment of the principles of the differential calculus, free of any consideration of infinitely small or vanishing quantities, of limits or of fluxions, and reduced to the algebraic analysis of finite quantities.<sup>34</sup>

If we wish to view Cauchy's new analysis in terms of structures, it seems clear that the new standards of proof it required not only changed the face but even the "look" of analysis. Cauchy's rigorous epsilontic calculus was perhaps just as revolutionary as the original discovery of the calculus by Newton and Leibniz had been.

---

cerned with the rigor of their arguments for pedagogical reasons – their interests were both more technical and more philosophical.

33 Grattan-Guinness, "Bolzano, Cauchy and the 'New Analysis' of the Early Nineteenth Century", *Archive for History of Exact Sciences*, 6 (1970), pp. 372–400; Grabiner 1981, pp. 11 and 74.

34 Grabiner 1981, pp. 138–139.



Again, as Grabiner has said:

It was not merely that Cauchy gave this or that definition, proved particular existence theorems, or even presented the first reasonably acceptable proof of the fundamental theorem of calculus. He brought all of these things together into a logically connected system of definitions, theorems, and proofs.<sup>35</sup>

In turn, the greater precision made possible by Cauchy's new foundations led to the discovery and application of concepts like uniform convergence and continuity, summability, and asymptotic expansions, none of which could be studied or even expressed in the conceptual framework of 18th-century mathematics. Names alone: Abel's convergence theorem, the Cauchy criterion, Riemann integral, Bolzano-Weierstrass theorem, Dedekind cut, Cantor sequences – all are consequences and reflections of the new analysis.

Moreover, there is again that important visual indicator of revolutions – a change in language reflected in the symbols so ubiquitously associated with the new calculus – namely deltas and epsilons – both of which first appear in Cauchy's lectures on the calculus in 1823.

In an extreme but telling example of the conceptual difference that separated Newton and Cauchy – at least when it came to conceiving of and justifying their respective versions of calculus – Grabiner tells the story of a student who asks what "speed" or "velocity" means, and is given an answer in terms of deltas and epsilons:

"The student might well respond in shock", she says, "How did anybody ever think of such an answer?"<sup>36</sup> The equally important question is "why" – why did Cauchy reformulate the calculus as he did? One answer, for greater clarity and rigor, seems obvious. By eliminating infinitesimals from polite conversation in calculus, and by substituting the arithmetic rigor of inequalities, he transformed a great part of mathematics, especially the language analysis would use and the standards to which its proofs would be held, for the next century and more. And yet, in the infinitesimals that Cauchy had so neatly avoided, lay the seeds of yet another, contemporary revolution in mathematics.

#### *4. Nonstandard Analysis as a Contemporary Revolution*

Historically, the dual concepts of infinitesimals and infinities have always been at the center of crises and foundations in mathematics, from the first "foundational crisis" that some, at least, have associated with discovery of irra-

---

<sup>35</sup> Grabiner 1981, p. 164.

<sup>36</sup> Grabiner 1981, p. 1.

tional numbers (or incommensurable magnitudes) by the Pythagoreans<sup>37</sup>, to the debates between 20th-century intuitionists and formalists – between the descendants of Kronecker and Brouwer on the one hand, and of Cantor and Hilbert on the other. Recently, a new "crisis" has been identified by the constructivist Errett Bishop:

There is a crisis in contemporary mathematics, and anybody who has not noticed it is being willfully blind. *The crisis is due to our neglect of philosophical issues...*<sup>38</sup>

Arguing that formalists mistakenly concentrate on "truth" rather than "meaning" in mathematics, Bishop has criticized nonstandard analysis as "formal finesse", adding that "it is difficult to believe that debasement of meaning could be carried so far".<sup>39</sup> Not all mathematicians, however, are prepared to agree that there is a crisis in modern mathematics, or that Robinson's work constitutes any debasement of meaning at all.

Kurt Gödel, for example, believed that Robinson, "more than anyone else", succeeded in bringing mathematics and logic together, and he praised Robinson's creation of nonstandard analysis for enlisting the techniques of modern logic to provide rigorous foundations for the calculus using *actual* infinitesimals. The new theory was first given wide publicity in 1961, when Robinson outlined the basic idea of his "nonstandard" analysis in a paper presented at a joint meeting of the American Mathematical Society and the Mathematical Association of America.<sup>40</sup> Subsequently, impressive applications of Robinson's

---

<sup>37</sup> There is a considerable literature on the subject of the supposed "crisis" in mathematics associated with the Pythagoreans, notably H. Hasse and H. Scholz, "Die Grundlagenkrise der griechischen Mathematik", *Kant-Studien*, 33 (1928), pp. 4–34. For a recent survey of this debate see J. L. Berggren, "History of Greek Mathematics: A Survey of Recent Research", *Historia Mathematica*, 11 (1984), pp. 394–410; Dauben 1984 [note 9 above]; D. H. Fowler, *The Mathematics of Plato's Academy: A New Reconstruction*, New York: Oxford UP, 1987; and W. Knorr, *The Evolution of the Euclidean Elements*, Dordrecht: Reidel, 1975.

<sup>38</sup> E. Bishop, "The Crisis in Contemporary Mathematics", *Proceedings of the American Academy Workshop in the Evolution of Modern Mathematics*, in *Historia Mathematica*, 2 (1975), p. 507. The emphasis is his.

<sup>39</sup> Bishop 1975, pp. 513–514.

<sup>40</sup> Robinson first published the idea of nonstandard analysis in a paper submitted to the Dutch Academy of Sciences. See A. Robinson, "Non-Standard Analysis", *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen*, ser. A, 64 (1961), pp. 432–440; reprinted in H. J. Keisler et al., eds., *Selected Papers of Abraham Robinson*, New Haven: Yale UP, 1979, vol. 2, pp. 3–11.

approach to infinitesimals have confirmed his hopes that nonstandard analysis could serve to enrich "standard" mathematics in substantive ways.

Using the tools of mathematical logic and model theory, Robinson succeeded in defining infinitesimals rigorously. He immediately saw this work not only in the tradition of others like Leibniz and Cauchy before him, but even as vindicating and justifying their views. The relation of their work, however, to Robinson's own research is equally significant (as Robinson himself realized), primarily for reasons that are of particular interest to the historian of mathematics.

This is not the place to rehearse the long history of infinitesimals. There is one historical figure, however, that especially interested Robinson, namely Cauchy, whose work provides a focus for considering the historiographic significance of Robinson's own work. In fact, following Robinson's lead, others like J. P. Cleave, Charles Edwards, Detlef Laugwitz and Wim Luxemburg have used nonstandard analysis to rehabilitate or "vindicate" earlier infinitesimalists.<sup>41</sup> Leibniz, Euler and Cauchy are among the more prominent mathematicians who have been "rationally reconstructed" – even to the point of having had, in the views of some commentators, "Robinsonian" nonstandard infinitesimals in mind from the beginning. The most detailed – and methodologically sophisticated of such treatments to date is that provided by Imré Lakatos.

##### 5. *Lakatos, Robinson and Nonstandard Interpretations of Cauchy's Infinitesimal Calculus*

In 1966 Imré Lakatos read a paper which provoked considerable discussion at the International Logic Colloquium meeting that year in Hannover. The primary aim of Lakatos' paper was made clear in its title: "Cauchy and the Continuum: the Significance of Non-standard Analysis for the History and

---

<sup>41</sup> J. P. Cleave, "Cauchy, Convergence and Continuity", *British Journal of the Philosophy of Science*, 22 (1971), pp. 27–37; C.H. Edwards 1979 [Note 10, above]; D. Laugwitz, "Zur Entwicklung der Mathematik des Infinitesimalen und Infinites", *Jahrbuch Überblicke Mathematik*, Mannheim: Bibliographisches Institut, 1975, pp. 45–50, and "Cauchy and Infinitesimals", *Preprint 911*, Darmstadt: Technische Hochschule Darmstadt, Fachbereich Mathematik, 1985; and W. A. J. Luxemburg, "Nichtstandard Zahlssysteme und die Begründung des Leibnizschen Infinitesimalkalküls", *Jahrbuch Überblicke Mathematik*, Mannheim: Bibliographisches Institut, 1975, pp. 31–44.

Philosophy of Mathematics".<sup>42</sup> Lakatos acknowledged his exchanges with Robinson on the subject of nonstandard analysis, which led to various revisions of the working draft of his paper. Although Lakatos never published the article, it enjoyed a rather wide private circulation and eventually appeared after Lakatos' death (in 1974) in volume 2 of his papers on "Mathematics, Science and Epistemology".

Lakatos realized that two important things had happened with the appearance of Robinson's new theory, indebted as it was to the results and techniques of modern mathematical logic. He took it above all as a sign that meta-mathematics was turning away from its original philosophical beginnings – and was growing into an important branch of mathematics.<sup>43</sup> This view, now more than twenty years later, seems fully justified.

The second claim Lakatos made, however, is that nonstandard analysis revolutionizes the historian's picture of the history of the calculus. The grounds for this assertion are less clear – and in fact, subject to question. In the words of Imré Lakatos:

Robinson's work...offers a rational reconstruction of the discredited infinitesimal theory which satisfies modern requirements of rigour and which is no weaker than Weierstrass's theory. This reconstruction makes infinitesimal theory an almost respectable ancestor of a fully-fledged, powerful modern theory, lifts it from the status of pre-scientific gibberish and renews interest in its partly forgotten, partly falsified history.<sup>44</sup>

Errett Bishop, somewhat earlier than Lakatos, was also concerned about the falsification of history, but for a different reason. Bishop explained the "crisis" he saw in contemporary mathematics in somewhat more dramatic terms:

---

<sup>42</sup> I. Lakatos, "Cauchy and the Continuum: The Significance of Non-standard Analysis for the History and Philosophy of Mathematics", in: J. Worrall and G. Currie, eds., *Mathematics, Science and Epistemology: Philosophical Papers*, vol. 2, Cambridge: Cambridge UP, 1978, pp. 148–151. Reprinted in *The Mathematical Intelligencer*, 1 (1979), pp. 151–161, with a note, "Introducing Imré Lakatos", pp. 148–151. Much of the argument developed here is drawn from a lengthier discussion of the historical and philosophical interest of nonstandard analysis in J. Dauben, "Abraham Robinson and Nonstandard Analysis: History, Philosophy and Foundations of Mathematics", in: P. Kitcher and W. Aspray, eds., *New Perspectives on the History and Philosophy of Mathematics*, Minneapolis: University of Minnesota Press, 1987, pp. 177–200; see as well "Abraham Robinson: Les Infinitésimaux, l'Analyse Non-Standard, et les Fondements des Mathématiques", in: H. Barreau, ed., *La Mathématique Non-Standard* (Fondements des Sciences), Paris: Editions du CNRS, 1989, pp. 157–184.

<sup>43</sup> Lakatos 1978, p. 43.

<sup>44</sup> Lakatos 1978, p. 44.

I think that it should be a fundamental concern to the historians that what they are doing is potentially dangerous. The superficial danger is that it will be and in fact has been systematically distorted in order to support the status quo. And there is a deeper danger: it is so easy to accept the problems that have historically been regarded as significant as actually being significant.<sup>45</sup>

Interestingly, Robinson sometimes made much the same point in his own historical writing. He was understandably concerned over the apparent triumph many historians (and mathematicians as well) have come to associate with the success of Cauchy-Weierstrassian epsilonotics over infinitesimals in making the calculus "rigorous". In fact, one of the most important achievements of Robinson's work has been his conclusive demonstration – thanks to nonstandard analysis – of the poverty of this kind of historicism. It is mathematically whiggish to insist upon an interpretation of the history of mathematics as one of increasing rigor over mathematically unjustifiable infinitesimals – the "*cholera baccillus*" of mathematics, to use Georg Cantor's colorful description of infinitesimals.<sup>46</sup>

Robinson, however, showed that there was nothing to fear from infinitesimals, and in this connection looked deeper, to the *structure* of mathematical theory, for further assurances:

Number systems, like hair styles, go in and out of fashion – it's what's underneath that counts.<sup>47</sup>

This might well be taken as the *leitmotiv* of much of Robinson's entire career, for his surpassing interest since the days of his dissertation written at the University of London in the late 1940's was model theory, and especially the ways in which mathematical logic could not only illuminate mathematics, but have very real and useful applications within virtually all of its branches. For Robinson, model theory was of such surpassing utility as a metamathematical tool because of its power and universality.

---

<sup>45</sup> Bishop 1975, p. 508.

<sup>46</sup> For Cantor's views, consult his letter to the Italian mathematician Vivanti, published in H. Meschkowski, "Aus den Briefbüchern Georg Cantors", *Archive for History of Exact Sciences*, 2 (1965), pp. 503–519, esp. p. 505. A general analysis of Cantor's interpretation of infinitesimals may be found in J. Dauben, *Georg Cantor: His Mathematics and Philosophy of the Infinite*, Cambridge, Mass.: Harvard UP, 1979; repr. Princeton: Princeton UP, 1990, pp. 128–132 and 233–238. On the question of rigor, refer to J. Grabiner, "Is Mathematical Truth Time-Dependent?", *American Mathematical Monthly*, 81 (1974), pp. 354–365.

<sup>47</sup> A. Robinson, "Numbers – What Are They and What Are They Good For?", *Yale Scientific Magazine*, 47 (1973), pp. 14–16.

In discussing number systems, Robinson wanted to demonstrate, as he put it, that:

The collection of all number systems is not a finished totality whose discovery was complete around 1600, or 1700, or 1800, but that it has been and still is a growing and changing area, sometimes absorbing new systems and sometimes discarding old ones, or relegating them to the attic.<sup>48</sup>

Robinson, of course, was leading up to the way in which nonstandard analysis had broken the bounds of the traditional Cantor-Dedekind understanding of the real numbers, just as Cantor and Dedekind had substantially transformed how continua were understood a century earlier in terms of Dedekind's "cuts", or even more radically with Cantor's theory of transfinite ordinal and cardinal numbers.<sup>49</sup>

There was an important lesson to be learned, Robinson believed, in the eventual acceptance of new ideas of number, despite their novelty or the controversies they might provoke. Ultimately, utilitarian realities could not be overlooked or ignored forever. With an eye on the future of nonstandard analysis, Robinson was impressed by the fate of another theory devised late in the 19th century which also attempted, like those of Hamilton, Cantor and Robinson, to develop and expand the frontiers of number.

In the 1890's Kurt Hensel introduced his now familiar p-adic numbers to investigate properties of the integers and other numbers. He also realized that the same results could be obtained in other ways. Consequently, many mathematicians came to regard Hensel's work as a pleasant game, but as Robinson himself observed, "many of Hensel's contemporaries were reluctant to acquire the techniques involved in handling the new numbers and thought they constituted an unnecessary burden".<sup>50</sup>

The same might be said of nonstandard analysis, particularly in light of Robinson's transfer principle that for any nonstandard proof in  $R^*$  (the extended nonstandard system of real numbers containing both infinitesimals and infinitely large numbers), there is a corresponding standard proof, complicated thought it may be. Moreover, many mathematicians are clearly reluctant to master the logical machinery of model theory with which Robinson developed his original version of nonstandard analysis. Thanks to Jerome Keisler and Wim Luxemburg, among others, nonstandard analysis is now accessible to mathematicians *without* their having to learn mathematical logic as a prerequi-

---

48 Robinson 1973, p. 14.

49 Dauben 1979 [note 46 above].

50 Robinson 1973, p. 16.

site.<sup>51</sup> For those who see nonstandard analysis as a fad, no more than a currently pleasant game like  $p$ -adic numbers, the later history of Hensel's ideas should give skeptics an example to ponder. Today,  $p$ -adic numbers are regarded as co-equal with the reals, and have proven a fertile area of mathematical research.

The same has been demonstrated by nonstandard analysis, for its applications in areas of analysis, the theory of complex variables, mathematical physics, economics, and a host of other fields have shown the utility of Robinson's own extension of the number concept. Like Hensel's  $p$ -adic numbers, nonstandard analysis can be avoided, although to do so may complicate proofs and render the basic features of an argument less intuitive.

What pleased Robinson about nonstandard analysis (as much as the interest it engendered from the beginning among mathematicians) was the way it demonstrated the indispensability, as well as the power, of technical logic:

It is interesting that a method which had been given up as untenable has at last turned out to be workable and that this development in a concrete branch of mathematics was brought about by the refined tools made available by modern mathematical logic.<sup>52</sup>

Robinson had begun his career as a mathematician by studying set theory and axiomatics with Abraham Fraenkel at Hebrew University in Jerusalem. Following his important work as an applied mathematician during World War II at the Royal Aircraft Establishment in Farnborough, he eventually went on to earn his Ph. D. from London University in 1949.<sup>53</sup> His early interest in logic was amply repaid in the applications he was able to make of logic and model theory first to algebra and somewhat later to the development of nonstandard analysis. As Simon Kochen has said of Robinson's contributions to mathematical logic and model theory:

---

<sup>51</sup> H. J. Keisler, *Elementary Calculus: An Approach Using Infinitesimals*, Boston: Prindle, Weber and Schmidt, 1976, and W. A. J. Luxemburg, *Lectures on A. Robinson's Theory of Infinitesimals and Infinitely Large Numbers*, Pasadena: California Institute of Technology, rev. ed., 1964.

<sup>52</sup> Robinson 1973, p. 16.

<sup>53</sup> Robinson completed his dissertation, "The Metamathematics of Algebraic Systems", at Birkbeck College, London University, in 1949. It was published two years later: *On the Metamathematics of Algebra*, Amsterdam: North-Holland Publishing Co., 1951. Several biographical accounts of Robinson are available, including G. Seligman, "Biography of Abraham Robinson", in: H.J. Keisler, et al., eds., *Selected Papers of Abraham Robinson*, New Haven: Yale UP, 1979, vol. 1, pp. xiii-xxxii; and J. Dauben, "Abraham Robinson", *The Dictionary of Scientific Biography, Supplement II* (New York: Scribners, 1990), pp. 748-751.

Robinson, via model theory, wedded logic to the mainstreams of mathematics.... At present, principally because of the work of Abraham Robinson, model theory is just that: a fully-fledged theory with manifold interrelations with the rest of mathematics.<sup>54</sup>

If the revolutionary character of nonstandard analysis is to be measured in textbook production and opposition to the theory, then it meets these criteria as well. The first textbook to teach the calculus using nonstandard analysis was written by Jerome Keisler, published in 1971, and opposition was expected. As G. R. Blackley warned Keisler's publishers (Prindle, Weber and Schmidt) in a letter when he was asked to review the new textbook prior to its publication:

Such problems as might arise with the book will be *political*. It is revolutionary. Revolutions are seldom welcomed by the established party, although revolutionaries often are.<sup>55</sup>

One member of the establishment who did greet Robinson's work with enthusiasm and high hopes was Kurt Gödel. It might even be said that he valued it for its "protean character", for it succeeded, he realized, in uniting mathematics and logic in an essential, fundamental way. That union has proven to be not only one of considerable mathematical importance, but of substantial philosophical and historical content as well.<sup>56</sup>

## 6. The Nature of Mathematical Resolution

In juxtaposing the concept of mathematical resolution with that of revolution, I have deliberately sought to make clear what I take to be the nature of scientific advance reflected in the development of the history of mathematics – including its most dramatic, revolutionary moments. Among these, as I have argued, are:

---

<sup>54</sup> S. Kochen, "Abraham Robinson: The Pure Mathematician. On Abraham Robinson's Work in Mathematical Logic", *Bulletin of the London Mathematical Society*, 8 (1976), pp. 312–315, esp. p. 313.

<sup>55</sup> K. Sullivan, "The Teaching of Elementary Calculus Using the Nonstandard Analysis Approach", *American Mathematical Monthly*, 83 (1976), pp. 370–375, esp. p. 375.

<sup>56</sup> On the "protean" character of mathematics, refer to the contribution to the San Sebastian Symposium published in this volume by Saunders MacLane, "The Protean Character of Mathematics". On Gödel and the high value he placed on Robinson's work as a logician, consult Kochen 1976, p. 316, and a letter from Kurt Gödel to Mrs. Abraham Robinson of May 10, 1974, quoted in Dauben 1990, p. 751.



1. The Greek discovery of incommensurable magnitudes (and the concomitant creation of a theory of proportion to accommodate them).
2. The Newton-Leibniz calculus.
3. The consequences of Cauchy's rigorization of analysis (through his epsilon-otics of "limit avoidance"<sup>57</sup>).
4. The development of transfinite set theory by Georg Cantor (arising from his profound discovery of the non-denumerability of the real numbers and his subsequent creation of transfinite numbers applicable to abstract set theory).
5. The creation of nonstandard analysis by Abraham Robinson in the 1960's.

In the spirit of the San Sebastian Symposium on Structures in Mathematical Theories, it should be added that each of these revolutions is intimately related to the structure of mathematics itself. Because the progress of mathematics is restricted only by the limits of self-consistency, the inherent structure of logic determines the structure of mathematical evolution. I have already suggested the way in which this evolution is necessarily cumulative. As theory develops, it provides more complete, more powerful, more comprehensive problem-solutions, sometimes yielding entirely new and revolutionary theories in the process. But the fundamental character of such advance is embodied in the idea of resolution. Like the microscopist, moving from lower to higher levels of resolution, successive generations of mathematicians can claim to understand more, with a greater stockpile of results and increasingly refined techniques at their disposal.

As mathematics becomes increasingly articulated, the process of resolution brings the areas of research and subjects of problem solving into greater focus, until solutions are obtained or new approaches developed to extend the boundaries of mathematical knowledge even further. Discoveries accumulate, and some inevitably lead to revolutionary new theories uniting entire branches of study, producing new points of view – sometimes wholly new disciplines – that would have been impossible to produce within the bounds of previous theory.

Revolutions in mathematics may not involve crisis or the rejection of earlier mathematics, although each of the revolutions I have discussed here represents a different response to the failures and limitations of prevailing theory. New discoveries, particularly those of revolutionary import, provide

---

<sup>57</sup> For detailed discussion of Cauchy and "limit avoidance", see I. Grattan-Guinness, *The Development of the Foundations of Mathematical Analysis from Euler to Riemann*, Cambridge, Mass.: MIT Press, 1970, esp. pp. 55–58, and pp. 72–74.

new modes of thought within which more powerful and general results are possible than ever before. To the question of whether or not revolutions occur in mathematics, my answer is an emphatic "yes".

## Observations, Problems and Conjectures in Number Theory – The History of the Prime Number Theorem

JAVIER ECHEVERRIA (San Sebastian)

### 1. Introduction

Mathematics has been considered by philosophers in the XXth century to be a non-experimental science. When Carnap proposed his division of Sciences in *Formalwissenschaften* (Logic, Mathematics...) and *Realwissenschaften* (Physics, Chemistry, Biology, Psychology, Sociology...) a strict epistemological and methodological demarcation was established, separating mathematics from the rest of the sciences.<sup>1</sup> Simultaneously, increasing Kantian influence introduced the doctrine that mathematical knowledge is *a priori*.<sup>2</sup>

More recently, W. V. Quine, H. Putnam, I. Lakatos and others<sup>3</sup> have criticized this *mathematical apriorism* in different ways,<sup>4</sup> suggesting that mathematics are also connected with empirical or quasi-empirical inferences and methods. The 1965 London Colloquium on "Empiricism in mathematics"<sup>5</sup>

---

1 R. Carnap, "Formalwissenschaften und Realwissenschaften", *Erkenntnis* 5 (1935), pp. 30–31.

2 Ph. Kitcher says that "Descartes, Locke, Berkeley, Kant, Frege, Hilbert, Brouwer and Carnap all developed the central apriorist thesis in different ways" (*The Nature of Mathematical Knowledge*, New York: Oxford UP 1984, p. 3).

3 See W. V. Quine, "Two Dogmas of Empiricism", in *From a Logical point of view*, chapter 7, New York: Harper & Row 1963; H. Putnam, "What is mathematical Truth?", in *Philosophical Papers*, vol. 1, Cambridge: Cambridge UP 1975; and I. Lakatos, *Proofs and Refutations*, Cambridge: Cambridge UP 1976.

4 See also W. Aspray and Ph. Kitcher (eds.), *History and Philosophy of Mathematics*, Minneapolis: Univ. of Minnesota Press 1988, for a history of critics of mathematical empiricism, specially their "An opinionated introduction", pp. 3–60.

5 I. Lakatos (ed.), *Problems in the Philosophy of Mathematics*, Amsterdam: North Holland 1967.

may be considered to be a point of no return for the *standard view* in the philosophy of mathematics. The present interest of philosophers in the history of mathematics, together with the revival of Polya's ideas<sup>6</sup> and the new conceptions which have been proposed in the most recent books about these topics,<sup>7</sup> suggests that something is changing in the philosophy of mathematics at the end of our century.

This paper will contribute in the same way by proposing some arguments, and a main example, arising out of Number Theory (NT). Despite its great importance from a methodological point of view, NT seems to be less attractive to philosophers than statistics, geometry or analysis. Consequently, we will try to draw attention to NT, called by Gauss the "Queen of mathematics".

## 2. Experimental Methods in Number Theory

Reading contemporary papers on NT, we find frequently similar assertions to Williams' claim in his article "Computers in Number Theory":

"Many number theorists perceive and practise their subject as an experimental science (...) In order to obtain some understanding of how things seem to be going in any particular *problem*, a number theorist will often compute *tables* of numbers related to the problem. Frequently some of the numbers in these tables exhibit a sort of pattern and this may indicate how a possible hypothesis concerning the problem can be developed. If further calculations confirm this hypothesis, the number theorist has good reason to believe that he may have a theorem within his grasp (...) Indeed, the numbers in these tables often provide some sort of idea as to how a proof should proceed. The speed and power of modern computers makes this program of research feasible, even when (as is frequent) many extensive and tedious calculations have to be performed".<sup>8</sup>

This kind of thesis is also very common in classics of NT, such as Euler, Gauss, Lucas and, more recently, Weil. The French mathematician Lucas, for example, says:

---

<sup>6</sup> See G. Polya, *Mathematics and Plausible Reasoning*, Princeton: Princeton UP 1954 (second ed. 1990), 2 vols., "Heuristic Reasoning in the Theory of Numbers", *Amer. Math. Monthly* 66 (1959), pp. 375–384 and the number 60:5(1987) of the *Mathematics Magazine*, which was published honoring Polya after his death.

<sup>7</sup> Kitcher 1984, *o.c.*; Aspray and Kitcher 1988, *o.c.*; Th. Tymoczko, *New directions in the Philosophy of Mathematics: an Antology*, Boston: Birkhäuser 1986; John Bigelow, *The Reality of Numbers: a Physicalist's Philosophy of Mathematics*, Oxford: Clarendon Press 1988.

<sup>8</sup> H.C. Williams, "The influence of computers in the development of Number Theory", *Comp. & Math. with Applications* 8:2 (1982), p. 75.

"Comme toutes les sciences, l'Arithmétique résulte de l'observation; elle progresse par l'étude de phénomènes numériques donnés par des calculs antérieurs, ou fabriqués, pour ainsi dire, par l'expérimentation; mais elle n'exige aucun laboratoire et possède seule le privilège de convertir ses inductions en théorèmes déductifs"... "C'est par l'observation du dernier chiffre dans les puissances des nombres entiers que Fermat, notre *Divus Arithmeticus*, créa l'Arithmétique supérieure, en donnant l'énoncé d'un théorème fondamental; c'est par la méthode expérimentale, en recherchant la démonstration de cette proposition, que la théorie des racines primitives fût imaginée par Euler; c'est par l'emploi immédiat de ces racines primitives que Gauss obtint son célèbre théorème sur la division de la circonférence en parties égales".<sup>9</sup>

André Weil, one of the most famous number theorists at the present time, calls for the existence of experiments in NT:

"Many people think that one great difference between Mathematics and Physics is that in Physics there are theoretical physicists and experimentalists and that a similar distinction does not occur in Mathematics. This is not true at all. In Mathematics just as in Physics the same distinction can be made, although it is not always so clear-cut. As in Physics the theoreticians think the experimentalists are there only to get the evidence for their theories while the experimentalists are just as firmly convinced that theoreticians exist only to supply them with nice topics for experiments. To experiment in mathematics means trying to deal with specific cases, sometimes numerical cases. For instance, an experiment may consist in verifying a statement like Goldbach's conjecture for all integers up to 1000, or (if you have a big computer) up to one hundred billion. In other words, an experiment consists in treating rigorously a number of special cases until this may be regarded as good evidence for a general statement" ... "Fermat was clearly a theoretician"... "Euler, on the other hand, was basically an experimentalist".<sup>10</sup>

If we believe number theorists themselves, simple induction, observation and experimentation have been frequently employed by mathematicians to propose problems and to state hypotheses and theorems about them, before finding rigorous methods of proof. This fact was noticed by some historians of mathematics, such as Scriba,<sup>11</sup> Goldstein<sup>12</sup> and Diamond,<sup>13</sup> setting out some

9 E. Lucas, *Théorie des Nombres*, Préface, p. XI, Paris: Blanchard 1958, second ed.

10 A. Weil, "Essais historiques sur la théorie des nombres", *L'enseignement mathématique* 21 (1975), pp. 14–15.

11 Ch. J. Scriba, "Zur Entwicklung der additiven Zahlentheorie von Fermat bis Jacobi", *Jber. Deutsch. Math.-Verein.* 72 (1970), pp. 122–142.

12 L.J. Goldstein, "A History of the Prime Number Theorem", *American Mathematical Monthly* 80 (1973), pp. 599–615.

13 Harold G. Diamond, "Elementary Methods in the Study of the Distribution of Prime Numbers", *Bull. of the Amer. Math. Soc.*, 7:3 (1982), pp. 553–589.

examples of discoveries in NT (Euler, Gauss, Jacobi, etc.) which were produced by observations of tables and consequent conjectures. In a more systematic way, Polya compared the research in NT with similar situations to those of physicists and chemists looking at the laws of Kepler, the Balmer's formula or discovering the Law of Multiple Proportions.<sup>14</sup> As a first conclusion about the methods of NT, we can state that the context of discovery is, in mathematics too, very different from the context of justification. Euler pointed this out in his article "De inductione ad plenam certitudinem evehenda":

"Notum est, plerasque numerorum proprietates primum per solam inductionem esse observatas, quas deinceps geometrae solidis demonstrationes confirmare elaboraverunt; in quo negotio im primis Fermatius summo studio et satis felici successu fuit occupatus".<sup>15</sup>

It would be easy to mention more quotations of the most important mathematicians which would certainly confirm the use of inductive, observational and empirical methods in NT. Instead of that, we can reconstruct the method of research of an ideal number theorist as follows:

*First*, an arithmetical question or problem is proposed to him by another mathematician, or is invented by the theorist himself.

*Second*, he observes several published tables of numbers or constructs his own tables and calculus, trying to find some way of solving the problem.

*Third*, he finds some meaningful facts concerning this problem and, frequently by simple induction, he proposes a hypothesis which could give it a first tentative solution. He confirms (or refutes) his hypothesis by contrasting it with new observations on tables or with concrete consequences of previously proven theorems. Of course, these "arithmetical facts" (or phenomena, as Lucas said) seem to be very different from physical or biological facts: I will come back to this question at the end of my paper.

*Fourth*, as a result of observations, trials, tests, comparisons and experiments with some contrasted arithmetic results, our number theorist is becoming convinced of the truth (or falsity) of a statement concerning the proposed problem. He can now publish his claim as a reasonable hypothesis in a mathematical journal. From an epistemic point of view, he knows for certain that

---

<sup>14</sup> Polya 1959, *o.c.*, p. 378.

<sup>15</sup> L. Euler, *Commentationes Arithmeticae collectae*, St. Petersburg 1849, Vol. 2, p. 134.

his hypothesis is true, because it is confirmed by relevant arithmetical facts. Despite its lack of proof, the proposition can be published together with the arguments and facts which give it corroborative evidence.

*Lastly*, the proposed problem and its hypothetical solution become a legitimate conjecture for the community of mathematicians. If it is interesting, everyone tries to confirm it by proof, or to refute it giving a counterexample or deducing from it a consequence which could contradict the accepted theory.

The history of NT is rich in such kind of *problems and conjectures* (we use both terms only when the proposed question has been received by the scientific community as a well-proposed problem or as a reasonable conjecture). Some of these hypothetical propositions will be proven, in a positive or in a negative way, but almost all of them can remain for a long time as genuine mathematical questions, contributing to the progress of mathematical theories. It is not necessary to agree with Polya's or Lakatos' theses to accept that some problems have been the main catalytic agent of the development of NT. Without mentioning the 10th Hilbert's problem, we could quote, for example, the books of Shanks and Guy<sup>16</sup> concerning solved and unsolved problems in NT, and similarly the sections proposing mathematical Problems in several important journals, such as the American Mathematical Monthly. Ingeniously, Guy proposed distinguishing between "good" and "bad" arithmetical problems, according to the time needed to solve them.<sup>17</sup> Among the good ones are, for example, Riemann's extended hypothesis (the 8th. Hilbert's Problem) and Goldbach's conjecture, which still remain unsolved, despite mathematician's general conviction that both are true. Undoubtedly, both conjectures have contributed more to the progress of Mathematics and Number Theory than many well-proved theorems.

We have distinguished five moments in which our idealized number theorist is working in the same way as an experimental scientist. After this, his results can be presented in a deductive way, even as an axiomatic theory. However, it would be a very strong simplification of NT to consider that only the context of justification must be analysed in order to reconstruct theories and methods of NT. Mathematical research is much more complicated than it seems if we are only readers of text books and commentators on the texts in which scientists try to explain their work. Philosophers of mathematics must

---

<sup>16</sup> D. Shanks, *Solved and Unsolved Problems in Number Theory*, New York: Chelsea 1978 (second edition); R. K. Guy: *Unsolved Problems in Number Theory*, New York: Springer 1981.

<sup>17</sup> R.K. Guy, *o.c.*, p. VII.

approach mathematical theories and methods in a more accurate way. History of mathematics provides sufficient material for this approach.

Consequently, I will try to go deeply into the history of NT, searching for case-studies which show us the way number theorists use empirical methods in their research. Instead of the usual speculation about the problematic empirical origin of mathematics, I will argue for and with careful studies of several key-moments of the history of NT. Empiricism in mathematics is, above all, a methodological and historical matter.

### 3. The Prime Number Theorem

From a methodological point of view, NT offers a very rich field for historical and philosophical studies. Reading a book such as Dickson's *History of the Theory of Numbers*,<sup>18</sup> we can easily find many meaningful examples, in which mathematicians use empirical methods: Fermat's theorem, Euler's law on sums of divisors of a number, Cunningham's Project, etc. It is my purpose to present the Prime Number Theorem (PNT) as the first case-study. There are several important reasons for this choice:

(a) PNT is the most relevant theorem concerning the distribution of prime numbers. Usually, it is considered to be, and named as, a *fundamental theorem* of arithmetics.

(b) PNT has been adequately studied by historians of mathematics. Just from 1896 to 1919, we can mention, for example, its exhaustive historical analysis by Landau,<sup>19</sup> the article by Hadamard,<sup>20</sup> the extensive report by Torelli,<sup>21</sup> the summaries by Landau,<sup>22</sup> Hardy and Littlewood<sup>23</sup> and, finally, the papers by Axer,<sup>24</sup> Landau again<sup>25</sup> and Steffensen.<sup>26</sup> More recently, Goldstein and Diamond summarized the history of PNT in the above mentioned articles.

(c) PNT was conjectured by Gauss in 1793, but it was not proved until

---

18 L. E. Dickson, *History of the Theory of Numbers*, Washington: Carnegie Institution 1920, 3 vols., rempr. New York: Chelsea, 1966.

19 E. Landau, *Handbuch der Lehre von der Verteilung der Primzahlen*, New York: Chelsea 1974 (1st. ed. 1909).

20 *Encyclopédie des sciences mathématiques* I, vol. 3, pp. 310–345.

21 Atti R. Accad. Sc. Fis. Mat., Napoli (2), 11(1902), No. 1, 222 pp.

22 *Proc. Fifth Inter. Congress of Math.*, Cambridge I (1913), pp. 93–108.

23 *Acta Mathematica* 41 (1917), pp. 119–196.

24 *Sitzungsber. Ak. Wiss. Wien (Math.)*, 120 (1911), IIa, pp. 1253–98.

25 *Ibid.*, 120 (1911), IIa, 973–88.

26 *Acta Mathematica* 37 (1914), pp. 75–112.



1896, when J. Hadamard and C. J. de la Vallée-Poussin published simultaneously independent proofs. According to Guy, it could be considered to be a "very good problem" for NT.

(d) After its first proof, mathematicians provided several different proofs of PNT, using different methods. At the beginning of this century, when the analytical methods for NT were increasing at the expense of elemental methods, several eminent mathematicians, such as Hardy,<sup>27</sup> Bohr<sup>28</sup> and Ingham,<sup>29</sup> stated that it would be almost impossible to find an elementary proof of PNT: only analytical methods, involving integral calculus and complex variables, could justify the truth of PNT. However, Erdős<sup>30</sup> and Selberg<sup>31</sup> produced two elementary and independent proofs of PNT in 1949, opening again the controversy about elementary and analytical methods.

Briefly, the historical and philosophical analysis and reconstruction of mathematical theories (rival theories?) associated with PNT will be the final aim of my paper, because of its great relevance within NT. It could be said that PNT represents a "paradigm" for mathematical research in NT.

#### 4. A Short History of PNT

Since Euclid, it is well known and proven that there are infinite prime numbers in  $\mathbb{N}$ . However, no general formula has ever been discovered to determine the prime number  $p_{n+1}$  from its previous  $p_n$ , nor has there been found any general law concerning the distribution of prime numbers within  $\mathbb{L}$ . Euler gave a good account of the last problem, which I will denote PNP (Prime Number Problem):

"Les mathématiciens ont tâché jusqu'ici en vain à découvrir un ordre quelconque dans la progression des nombres premiers, et on a lieu à croire, que c'est

---

<sup>27</sup> G.H. Hardy, "Prime Numbers", *British Assoc. Reports* 1915, pp. 350–354, repr. in *Collected Papers*, vol. 2, London: Oxford UP 1967, pp. 14–18.

<sup>28</sup> See H. Bohr and H. Cramer, "Die neuere Entwicklung der analytischen Zahlentheorie", *Enzyklopädie der mathematischen Wissenschaften* II c 8 (1922), pp. 722–849. See also *Proc. Int. Congress of Math.*, vol. I (1952), pp. 127–132.

<sup>29</sup> A.E. Ingham, *The distribution of Prime Numbers*, Cambridge: Cambridge UP 1932, 2d ed. 1990.

<sup>30</sup> P. Erdős, "On a new method in elementary number theory which leads to an elementary proof of the prime number theorem", *Proc. Nat. Acad. Sci. USA*. 35 (1949), pp. 374–384.

<sup>31</sup> A. Selberg, "An elementary Proof of the Prime Number Theorem", *Ann. of Math.*, 50:2 (1949), pp. 305–313.

un mystère auquel l'esprit humain ne saurait jamais pénétrer. Pour s'en convaincre, on n'a qu'à jeter les yeux sur les tables des nombres premiers, que quelques personnes se sont données la peine de construire au-delà de cent mille: et on s'apercevra d'abord qu'il n'y regne aucun ordre ni règle. Cette circonstance est d'autant plus surprenante, que l'arithmétique nous fournit des règles sûres, par le moyen desquelles on est en état de continuer la progression de ces nombres aussi loin que l'on souhaite, sans pourtant nous y laisser apercevoir la moindre marque d'un ordre quelconque. Je me vois aussi bien éloigné de ce but, mais je viens de découvrir une loi fort bizarre parmi les sommes des diviseurs des nombres naturels, sommes qui, au premier coup d'oeil, paraissent aussi irrégulières que la progression des nombres premiers, et qui semblent même envelopper celle-ci. Cette règle, que je vais expliquer, est à mon avis d'autant plus importante qu'elle appartient à ce genre de vérités dont nous pouvons nous persuader, sans en donner une démonstration parfaite. Néanmoins, j'en allèguerai des preuves telles, qu'on pourra presque les envisager comme équivalentes à une démonstration rigoureuse".<sup>32</sup>

In the same article, Euler employed empirical methods (observation, induction) to find an unexpected law which provided a first solution to an arithmetical problem, the sum of the divisors of a number, which could seem before Euler so untractable and irregular such as the distribution of prime numbers (PNP). After Euler's revival of empirical methods in arithmetics, and mainly after Lambert's strong appeals in 1770 to construct factor tables and prime number tables,<sup>33</sup> the most important mathematicians frequently began to work with extensive prime number tables (PNt), trying to find in them some regularity which could yield properties or hypotheses about these kinds of arithmetical problems.

The first real advances in PNP were made by Gauss (1793) and Legendre (1798). The second was the first author to publish a hypothesis about PNP<sup>34</sup>:

If we denote  $\pi(x)$  the number of prime numbers  $p$  less than  $x$  ( $p < x$ ,  $x \in \mathbb{N}$ ,  $p$  prime number), Legendre's hypothesis of 1798 stated the similarity between

$$\pi(x)$$

and

$$\frac{x}{A \log x + B}$$

<sup>32</sup> L. Euler, *Opera Arithmetica, Additamenta*, p. 639.

<sup>33</sup> J. H. Lambert, *Zusätze zu den logarithmischen und trig. Tabellen*, Berlin, 1770.

<sup>34</sup> A. M. Legendre, *Essai sur la Théorie des Nombres*, second ed., Paris 1808, part IV, 8.

A and B being parameters. Legendre improved his hypothesis in 1808, after the publication of the Vega's Prime Numbers Tables, and expressed it in the following way:

$$\pi(x) = \frac{x}{\log x + A}$$

where A was calculated by Legendre as approximately 1.80366. It seems that Legendre deduced that the average interval between two primes at the point x is of the form  $A \log x + B$ , and consequently that the number of primes inferior to x is approximately equal to

$$\frac{x}{A \cdot \log x + B - A}$$

"ce qui s'accorde avec la formule générale donnée ci-dessus, en prenant  $A=1$ ,  $B=-0.08366$ ".<sup>35</sup>

How the value of both constants were determined? James Glaisher suggested that:

"although no doubt the constant was determined mainly from  $x=10.000$ , it does not appear to have been derived from any single value of x; but it seems likely that it was so chosen as to represent as nearly as possible the results of the entire enumerations".<sup>36</sup>

In the Chapter entitled "D'une loi très remarquable observée dans l'énumération des nombres premiers" Legendre gave the Table PNT-3, (see below), in which the values obtained from the formula are compared with the numbers actually counted up to 400.000. After the Table he remarked:

"Il est impossible qu'une formule représente plus fidèlement une série de nombres d'une aussi grande étendue et sujette nécessairement à de fréquentes anomalies".<sup>37</sup>

---

<sup>35</sup> *Ibid.*, end of chap. 8.

<sup>36</sup> J. Glaisher, *Factor Table for the sixth million*, London: Taylor and Francis 1883, p. 67.

<sup>37</sup> A. M. Legendre, *o.c.*, cap. 8.

TABLE PNT-3

x	Formula	Tables
10.000	1.230	1.230
20.000	2.268	2.263
30.000	3.252	3.246
40.000	4.205	4.204
50.000	5.136	5.134
60.000	6.049	6.058
70.000	6.949	6.936
80.000	7.838	7.837
90.000	8.717	8.713
100.000	9.588	9.592
150.000	13.844	13.849
200.000	17.982	17.984
250.000	22.035	22.045
300.000	26.023	25.998
350.000	29.961	29.977
400.000	33.854	33.861
.....	.....	.....
400.000	33.854	33.863
500.000	41.533	41.538
600.000	49.096	49.093
700.000	56.565	56.535
800.000	63.955	63.937
900.000	71.279	71.268
1.000.000	78.543	78.493

In 1816 Legendre published a supplement to his *Théorie des Nombres*, which contains the continuation of the Table PNT-3 from 400.000 to 1.000.000 (see Table PNT-3). The new enumerations were made from Chernac's *Cribrum Arithmeticum*, published in 1811, which gives all the prime factors of numbers up to 1.020.000. About the new Table James Glaisher comments:

"It is remarkable that a constant which might have been determined from  $x=10.000$  should so accurately represent the numbers of primes up to 1000000".<sup>38</sup>

Gauss did not publish his own research on PNP, but his posthumous manus-

---

<sup>38</sup> J. Glaisher, *o.c.*, p. 68.

cripts<sup>39</sup> prove that he employed empirical and inductive methods, which were founded on the observation of prime number tables (PNT). For example, he constructed the table:

TABLE PNT - 1

Intervals	Primes	Intervals	Primes
1-1000	168	25000-26000	98
1000-2000	135	26000-27000	101
2000-3000	127	27000-28000	94
3000-4000	120	28000-29000	98
4000-5000	119	29000-30000	92
5000-6000	114	30000-31000	95
6000-7000	117	31000-32000	92
7000-8000	107	32000-33000	106
8000-9000	110	33000-34000	100
9000-10000	112	34000-35000	94
10000-11000	106	35000-36000	92
11000-12000	103	36000-37000	99
12000-13000	109	37000-38000	94
13000-14000	105	38000-39000	90
14000-15000	102	39000-40000	96
15000-16000	108	40000-41000	88
16000-17000	98	41000-42000	101
17000-18000	104	42000-43000	102
18000-19000	94	43000-44000	85
19000-20000	102	44000-45000	96
20000-21000	98	45000-46000	86
21000-22000	104	46000-47000	90
22000-23000	100	47000-48000	95
23000-24000	104	48000-49000	89
24000-25000	94	49000-50000	98

which suggested to him a first hypothesis about PNP:

PNH-1: If  $n \in \mathbb{N}$  is recovered by an interval  $(a,b)$  included in  $\mathbb{N}$ ,  $\pi(n)$  equals approximately to  $\int_a^b \frac{dx}{\log x}$ .

<sup>39</sup> C. F. Gauss, "Tafel der Frequenz der Primzahlen". In: *Werke*, vol. II, Göttingen 1863, pp. 435-447.

It is important to note that Gauss, trying to solve PNP, put forward the hypothesis PNH-1 suggested by tables PNT-1 and, in this way, he did not have problems in using the Integral Calculus. It is also very important to note that, in those years and following Lambert's claim, relevant advances were produced by Vega (1796) in order to construct more extensive prime number tables.<sup>40</sup> Gauss knew those tables, such as the posterior PNT of Chernac (1811) and Burckhardt (1817).<sup>41</sup>

Undoubtedly, the correspondence between Gauss and Encke must also be carefully considered.<sup>42</sup> In his letter to Encke dated the 24th November 1849, he wrote:

"Your remarks concerning the frequency of primes were of interest to me in more ways than one. You have reminded me of my own endeavors in this field which began in the very distant past, in 1792 or 1793, after I had acquired the Lambert supplements to the logarithmic tables. Even before I had begun my more detailed investigations into higher arithmetic, one of my first projects was to turn my attention to the decreasing frequency of primes, to which end I counted the primes in several chiliads and recorded the results on the attached white pages. I soon recognized that behind all of its fluctuations, this frequency is on the average inversely proportional to the logarithm, so that the number of primes below a given bound  $n$  is approximately equal to

$$\int \frac{dn}{\log n},$$

where the logarithm is understood to be hyperbolic. Later on, when I became acquainted with the list in Vega's tables (1796) going up to 400031, I extended my computation further, confirming that estimate. In 1811, the appearance of Chernac's *Cribrum* gave me much pleasure and I have frequently (since I lack the patience for a continuous count) spent an idle quarter of an hour to count another chiliad here and there; although I eventually gave it up without quite getting through a million. Only some time later did I make use of the diligence of Goldschmidt to fill some of the remaining gaps in the first million and to continue the computation according to Burckhardt's tables. Thus (for many years now) the first three million have been counted and checked against the integral".<sup>43</sup>

Knowing that Encke had proposed his own hypothesis about PNP,

PNH-2:  $\pi(x)$  is approximately equal to  $\frac{n}{\log n - 1.513}$ ,<sup>44</sup>

<sup>40</sup> G. Vega, *Tabulae logarithmico-trigonometricae*, 1797, vol. 2.

<sup>41</sup> L. Chernac, *Cribrum Arithmeticum*, Daventriae 1811.

<sup>42</sup> J. C. Burckhardt, *Table des diviseurs*, Paris 1814–1817.

<sup>43</sup> Gauss, *o.c.*, translation offered by L. J. Goldstein, *o.c.*, pp. 612–614.

<sup>44</sup> *Ibid.*

TABLE PNt-2

Below	Here are Prime	Integral $\int \frac{dn}{\log n}$	Error	Your Formula	Error
500000	41556	41606,4	+ 50,4	41596,9	+ 40,9
1000000	78501	79627,5	+ 126,5	78672,7	+ 171,7
1500000	114112	114263,1	+ 151,1	114374,0	+ 264,0
2000000	148883	149054,8	+ 171,8	149233,0	+ 350,0
2500000	183016	183245,0	+ 229,0	183495,1	+ 563,5
3000000	216745	216970,6	+ 225,6	217308,5	+ 563,5

This Table contains larger results than PNt-1, because of the progress achieved by Vega, Chernac and Burckhardt. In the same letter Gauss talked about Legendre's research on PNP:

"I was not aware that Legendre had also worked on this subject; your letter caused me to look in his *Théorie des Nombres*, and in the second edition I found a few pages on the subject which I must have previously overlooked (or, by now, forgotten). Legendre used the formula:

$$\frac{n}{\log n - A}$$

where A is a constant which he sets equal to 1.08366. After a hasty computation, I find in the above cases the deviations -23,3; +42,2; +68,1; +92,8; +159,1; +167,6. These differences are even smaller than those from the integral, but they seem to grow faster with n so that it is quite possible they may surpass them. To make the count and the formula agree, one would have to use, respectively, instead of A=1.08366, the following numbers: 1,09040; 1,07682; 1,07582; 1,07529; 1,07179; 1,07297. It appears that, with increasing n, the (average) value of A decreases; however, I dare not conjecture whether the limit as n approaches infinity is 1 or a number different from 1. I cannot say there is any justification for expecting a very simple limiting value; on the other hand, the excess of A over 1 might well be a quantity of the order of 1/logn. I would be inclined to believe that the differential of the function must be simpler than the function itself".<sup>45</sup>

If we compare the three proposed hypotheses concerning PNP, PNt-1 (Gauss), PNt-2 (Legendre) and PNt-3 (Encke), the first should take preference, according to the empirical criterion of Gauss, because it has been corroborated in a higher degree by the tables (PNt-1, PNt-2, etc.) which were available at that time. As a result, we can conclude that PNt-1 was the best possible solution in 1850, in comparison to several rival hypotheses and, of

<sup>45</sup> *Ibid.*

course, relative to the "arithmetic facts" which were provided by another class of mathematicians, *the computers*, following a complementary line of research: improvement of PNT. Consequently, we can also conclude that mathematical approaches to PNP were clearly empirical in Gauss' time, according to the norms of experimental methods. After the publication of Gauss' works, the community of mathematicians began to study PNP systematically.

We can point to this moment as the real origin of PNP, from an historical, sociological and philosophical point of view. The second phase began in 1849, when Fuss edited the arithmetical memoirs of Euler in two volumes, containing memoirs such as *De tabula numerorum primorum* or *De numeris primis valde magnis*, in which Euler compares PNP in Arithmetic to the quadrature of the circle in Geometry:

"Vix ullus reperiatur geometra, qui non, ordinem numerorum primorum investigando, haud parum temporis inutiliter consumserit: videtur enim lex, qua numeri primi progrediuntur, in arithmetica aequae abstrusae esse indaginis, atque in Geometria circuli quadratura: ac si hujus indagatio pro desperata est habenda, non leviores adsunt rationes, quae et ordinis, quo numeri primi se invicem sequuntur, cognitionem nos in perpetuum fugere persuadent".<sup>46</sup>

In July 1849 the *Philosophical Magazine* published a paper by C. J. Hargreave, entitled "Analytical Researches Concerning Numbers", in which the connection between the number of primes inferior to any given limit and the logarithm-integral was stated. Independently, Chebychev rediscovered the hypothesis of Gauss in his memoir "Sur la fonction qui détermine la totalité des nombres premiers inférieurs à une limite donnée": it appeared in the sixth volume of the Memoirs of the Academy at St. Petersburg (1851) and was reprinted in vol. XVII of Liouville's Journal (1852).<sup>47</sup> Consequently, the function

$$\lim(x) = \int_0^x \frac{dx}{\log x}$$

was employed by Gauss, Chebychev and Hargreave to approximately represent the number of primes inferior to  $x$ .

However, the most important research concerning the number of primes inferior to  $x$ ,  $\phi(x)$ , was produced by Riemann in 1859.<sup>48</sup> He showed that  $\phi(x)$  consists of a non-periodic function and a series of periodic terms involving the

<sup>46</sup> L. Euler, *o.c.*, section VI, pp. 36–38.

<sup>47</sup> P. Chebychev, "Mémoire sur les nombres premiers", *Mém. de l'Ac. de St. Petersburg (Savans Etrangers)*, VII (1854), pp. 15–33; *Journal of Liouville* XVII (1852), pp. 366–390.

<sup>48</sup> G. F. B. Riemann, "Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse", *Monatsber. d. Ak. d. Wiss. z. Berlin* 1859 (1860). pp. 671–680.



roots of a certain transcendental equation. The non-periodic function is

$$\lim(x) - 1/2.\lim(x^{1/2}) - 1/3.\lim(x^{1/3}) - 1/5.\lim(x^{1/5}) + 1/6.\lim(x^{1/6}) - \dots$$

where the terms are of the form  $\pm 1/n.\lim(x^{1/n})$ ,  $n$  being a product of prime factors of the form  $a \cdot b \cdot c$ , so that  $n$  contains no squared factor. Consequently, there appeared a new hypothesis concerning the number of primes inferior to  $x$ , PNH-4, quite different from PNH-1, 2 and 3 (we will identify the hypotheses of Gauss, Hargreave and Chebychev).

Simultaneously, Zacharias Dase improved PNt (1862–65) and published tables of prime numbers and factors from 6.000.000 to 9 millions.<sup>49</sup> The Committee on Mathematical Tables of the British Association for the Advancement of Sciences, composed by Cayley, Stokes, Smith, Thomson, J. Glaisher and J. W. L. Glaisher, considered the completion of the factor tables up to 10.000.000 a matter of great scientific importance and the Association accorded a grant for this purpose. From 1879 to 1885 three volumes were published by the Association, containing PNt for the 4th, 5th and 6th million.

We can consider this second phase of the history of PNP as the most important for our purpose. If we read Section III ("Comparison of the results of the enumeration with Legendre's, Chebychev's and Riemann's formulae") of Glaisher's *Factor Tables for the sixth million* (1883), we can distinguish five different Prime Number Hypotheses (PNH) to be contrasted with the extended Prime Number Tables (PNt): PNH-1 (Gauss-Hargreave-Chebychev), PNH-2 (Legendre), PNH-4 (Riemann) and two new hypotheses introduced by Glaisher by means of two analytic functions:

$$\text{PNH-5: } \frac{x}{\log x - 1 - \frac{1}{\log x}}$$

$$\text{PNH-6: } \frac{x}{\log x - 1}$$

Table VI of Glaisher's book shows us the deviations of the five hypotheses: for small values of  $x$ , PNH-3 is more accurate than PNH-1; both hypotheses predict the same value at  $x=4.850.000$ ; beyond this point Legendre's formula steadily diverges from the real values observed in PNt. Riemann's formula is always the most exact (see Table VI).

The last phase of the history of PNP began with two papers published by Meissel in the *Mathematische Annalen* (1870 and 1871), introducing a method to determine the number of primes inferior to  $x$  without an actual computation of PNt. Piarron de Mondesir proposed a different method in the *Comptes Ren-*

---

<sup>49</sup> Z. Dase, *Factoren-Tafeln für alle Zahlen...*, Hamburg 1862–1865, 3 vols.

TABLE VI  
DEVIATIONS OF THE FIVE FORMULÆ

x	(i) Riemann	(ii) Cheby- chev	(iii) Legendre	(iv) $\frac{x}{\log x - 1 - \frac{1}{\log x}}$	(v) $\frac{x}{\log x - 1}$
100000	-6	+37	-5	-2	+81
200000	-3	51	-3	-6	138
300000	+26	89	+26	+16	162
400000	-9	62	-7	25	245
500000	-9	67	-6	32	293
600000	-8	74	-3	37	337
700000	+13	101	+21	22	359
800000	-7	85	+3	48	421
900000	-9	87	+4	57	465
1000000	+29	129	+44	25	469
1100000	+22	126	+41	38	515
1200000	-41	67	-19	107	617
1300000	-2	109	+24	74	616
1400000	-24	90	5	101	676
1500000	-6	111	27	89	696
.....	.....	.....	.....	.....	.....
2300000	-20	119	45	147	994
2400000	+4	145	73	129	1005
2500000	+29	172	102	109	1014
2600000	-4	142	74	147	1080
2700000	-10	138	73	158	1119
2800000	+17	167	104	136	1126
2900000	-3	149	88	162	1179
3000000	-1	154	96	165	1209
3100000	+19	175	120	150	1222
3200000	-17	141	89	191	1290
3300000	-41	120	70	220	1345
3400000	-26	137	89	210	1362
3500000	+4	168	124	185	1364
3600000	+58	224	182	137	1342
3700000	+3	171	132	196	1429
3800000	+17	186	152	187	1445
3900000	-14	158	126	223	1508
4000000	+33	206	177	181	1492
4100000	-24	151	125	243	1579
4200000	-3	173	151	227	1589
4300000	+37	215	196	192	1580
4400000	+67	+247	+231	-167	-1581

TABLE VI (continued)

x	(i) Riemann	(ii) Cheby- chev	(iii) Legendre	(iv)	(v)
				$\frac{x}{\log x - 1 - \frac{1}{\log x}}$	$\frac{x}{\log x - 1}$
4500000	-16	+166	+154	-254	-1693
4600000	+13	196	187	232	1695
4700000	+2	186	181	246	1736
4800000	+30	216	214	223	1738
4900000	-30	158	160	287	1827
5000000	-66	123	129	329	1893
5100000	-47	144	153	314	1904
5200000	-14	179	192	285	1900
5300000	-46	148	165	322	1961
5400000	-55	141	161	335	1999
5500000	-47	150	174	333	2021
5600000	-11	187	215	301	2014
5700000	+16	215	247	279	2016
.....	.....	.....	.....	.....	.....
6700000	+55	268	340	286	2262
6800000	+38	252	329	307	2307
6900000	-21	194	275	371	2394
7000000	-40	177	262	394	2441
7100000	-69	150	239	426	2497
7200000	-40	180	273	403	2496
7300000	-59	162	260	426	2543
7400000	-34	188	291	405	2545
7500000	-16	207	314	392	2555
7600000	-8	216	327	389	2575
7700000	-47	178	294	432	2641
7800000	+10	237	357	378	2611
7900000	-56	172	296	450	2705
8000000	-37	192	320	435	2713
8100000	0	231	364	402	2703
8200000	-18	214	351	424	2748
8300000	-23	210	352	434	2780
8400000	-35	199	346	450	2819
8500000	-11	224	375	430	2822
8600000	-73	163	319	495	2911
8700000	-95	142	303	522	2959
8800000	-139	100	264	571	3030
8900000	-108	131	301	545	3026
9000000	-132	+108	+282	-573	-3077

due of the 'Association Française pour l'Avancement des Sciences'.<sup>50</sup> As a result, Meissel calculated the number of primes inferior to 100.000.000. Although simple in theory, the new methods for the calculations were quite difficult in application: the problem of the computational complexity of an algorithm appeared immediately.

From a theoretical point of view, important progress was made following the way opened by Chebychev and Riemann. The Russian mathematician introduced two new functions of a real variable  $x$ :

$$\Theta(x) = \sum \log p, \text{ being } p \text{ prime and } p \leq x,$$

and

$$\Psi(x) = \sum \log p, \text{ being } p \text{ prime and } p^m \leq x.$$

Chebychev proved that the Prime Number Problem (PNP) could be solved if we prove:

$$\lim_{x \rightarrow \infty} \frac{\Theta(x)}{x} = 1$$

or

$$\lim_{x \rightarrow \infty} \frac{\Psi(x)}{x} = 1$$

He proved also that if these limits exist, then their value must be 1. However, "Chebychev's methods were of an elementary, combinatorial nature, and – according to Goldstein's comment<sup>51</sup> – as such were not powerful enough to prove the prime number theorem".

Riemann proposed a new idea, employing the Riemann zeta function,

$$\zeta(s) = \sum_1^\infty \frac{1}{n \cdot n \cdots n}, \text{ where } s \text{ is a complex variable.}$$

Furthermore, he connected his *zeta function* and the problem of the distribution of primes, giving the *Riemann's explicit formula*:

$$\zeta(x) = x - \sum_p \frac{x^p}{p} - \frac{\zeta'(0)}{\zeta(0)} - \frac{1}{2} \log(1-x^{-2}),$$

<sup>50</sup> See *Math. Ann.*, II(1870), pp. 636–642 and III(1871), pp. 523–525; Piaron de Mondesir, "Formules pour le calcul exact de la totalité des nombres premiers...", *Compte rendu de l'Assoc. Fr. pour l'avancement des sciences*, Le Havre 1877–1878, pp. 79–92.

<sup>51</sup> L. Goldstein, *o.c.*, p. 606.

where  $r$  runs over all non-trivial zeroes of the Riemann *zeta function*.

Following Riemann's method and employing the most powerful instruments of the mathematical analysis, J. Hadamard and Ch. de la Vallée Poussin proved independently the Prime Number Theorem in 1896. The third phase of the history of PNP was closed with a great success for the community of mathematicians: the use of logarithms, series, integral calculus and functions of real and complex variable provided a solution to PNP, one of the most difficult problems in arithmetic.

However, this kind of success proving PNT is not so important for our purpose. The history of PNP has continued during the XXth century. At the beginning, analytical methods produced new results and theorems. In 1949 Erdős and Selberg proved independently PNT by elementary methods, that is, without using complex variable nor integral calculus: as a consequence, a strong revival of the Elementary Number Theory took place.<sup>52</sup> After the unexpected success of Erdős and Selberg, many different proofs of PNT have been published. In order to limit the field of research, I will limit my comments to the first and second phase of the history of PNP, where empirical methods were frequently used.

### 5. Different Proofs of PNT

If we summarize some of the most important moments of the history of PNP, we should especially emphasize the following points:

(a) Euler's empirical solution of a similar problem (the sum of the divisors of a number) introduced new methods in NT, such as simple induction, construction and observation of tables, hypotheses founded on empirical evidence, etc. As a consequence of these methods, Gauss and Legendre studied different tables in order to search for a solution of PNP.

(b) Gauss, Legendre and Encke proposed three rival hypotheses about PNP, which we have denoted PNH-1, PNH-2 and PNH-3. At that time, these hypotheses had only empirical and approximative confirmation. Prime number tables were simultaneously improved, making possible the empirical evaluation of the proposed hypotheses.

(c) Chebyshev rediscovered PNH-1 and introduced a new and fundamental function in NT:  $\text{Li}(x)$ . It was probably the first occasion that number theorists accepted integral methods to solve an arithmetical problem. His proof of Bertrand's Postulate was undoubtedly a great success for this kind of methods.

---

<sup>52</sup> See H.G. Diamond, *o.c.*

(d) Making use of an extended analytical method, Riemann proposed his new function, the *zeta function*. Consequently, the use of Complex Analysis together with Integral Calculus became very frequent among number theorists. Riemann stated a new hypothesis about PNP, PNH-4, connected to the more general *Riemann Hypothesis*.

(e) Chebychev defined two new elementary functions and introduced new and very simple expressions for PNP. Both functions had been the source of a new mathematical theory: the *theory of arithmetical functions*. Consequently, it can be concluded that several mathematical theories emerged from PNP before the proof of PNT. Progress in mathematics is sometimes produced by problems and conjectures, more than by theorems.

(f) Meissel discovered a more general proposition, the Meissel's Identity, very useful to solve different questions in NT. As a result, *recursive functions* were also admitted as a possible way of solving PNP. By means of his computational methods, Meissel simplified considerably the arithmetical calculations of  $\pi(x)$ . However, an unexpected obstacle arose: the efficiency of an algorithm is not always guaranteed. NT involves the search for faster and more efficient algorithms.

(g) Simultaneously, PNT were considerably extended and improved by different mathematicians devoted to computing. Vega, Chernac, Burckhardt, Dase, Glaisher and Meissel himself constructed tables of prime numbers up to 10 millions. This kind of research was supported by Scientific Societies. The problem of the storage of such Tables was also explicitly considered by Lebesgue and many others. In any case, a PNT became an indispensable tool for every number theorist.

(h) Therefore, it can be said that "empirical basis" of PNP became larger and more accurate during the XIXth century. The available tables at the end of that century made it possible to reject several proposed hypothesis (PNH-2 and PNH-3, for example), as well as a careful corroboration of other ones (PNH-1, PNH-4, etc.). Before the proof of PNT, empirical progress had decided the old controversy in favour of Gauss, Chebychev and Riemann's claims.

(i) From a theoretical point of view, however, the most promising method of research was proposed by Riemann and followed by many mathematicians. As a final result, J. Hadamard and Ch. de la Vallée-Poussin obtained the first proof of PNT with the aid of powerful analytical methods. At this time, supporters of elementary methods in NT were literally demolished. It could be said that a scientific revolution took place in Number Theory.

(j) The proof of PNT was followed by a real process of scientific change in Mathematics. Landau, Littlewood, Hardy, Ramanujan and many other mathe-

maticians successfully applied analytical methods to NT and discovered several relevant theorems, which were proved according to the most rigorous criteria of mathematical analysis.

(k) At the beginning of the XXth century different proofs of PNT were available to mathematicians and each one was developed in a very different theoretical framework, within analytical methods. In spite of that, defenders of elementary methods did not give up trying to find another kind of proof. Brun (1920), Titshmarsh and above all Schnirelman<sup>53</sup> produced important theorems using only elementary methods. Finally, Erdős and Selberg proposed a new formulation of an old Chebychev's formula and, employing it, found the first elementary proof of PNT. This unexpected success, of course, was followed by a powerful renaissance of the Elementary Number Theory: different versions of Selberg's formula appeared, PNT was generalized to arithmetic progressions and algebraic fields, abstract systems of numbers emerged as a consequence, etc. We can conclude that, from a heuristic point of view, different proofs of the same mathematical theorem produce very different consequences and theories.

## 6. The Empirical Basis of Number Theory

The main purpose of my present contribution states that prime number tables (PNT) play the same epistemological role in NT as empirical measures and data in experimental sciences. The Prime Number Theorem can easily be considered as a very illustrative example for this claim. We can argue several reasons to support this view:

(a) After the methodological change introduced by Euler in NT, construction of more extensive prime number tables became one of the most important aims for number theorists. Gauss's hypothesis was a direct result of Lambert's *Supplementa* (1792), which contains a list of primes up to 102.000. When Vega's *Tabulae* (1796) were published, Gauss was able to extend his enumeration of primes up to 400.031. In 1811 Chernac's *Cribrum Arithmeticum* appeared and Gauss enlarged immediately his own tables and eventually confirmed his hypothesis for higher values, up to 3.000.000. The editor of Gauss's *Werke*, Dr. Schering, states that the Table of the first million was handwritten by Gauss,<sup>54</sup> who was helped by Goldschmidt in making PNT from

<sup>53</sup> V. Brun, "La série  $1/5+1/7+1/11+1/13+\dots$  où les dénominateurs sont nombres premiers jumeaux est convergente ou finie", *Bull. Sci. Math.* (2) 43 (1919), pp. 100–104 and 124–128.

<sup>54</sup> Gauss, *o.c.*, ed. 1876, vol. 2, p. 521.

1.000.000 to 3.000.000. The firm conviction of Gauss about the truthlikeness of his hypothesis depended on these successive empirical corroborations. Briefly, we can consider this hypothesis to be a *prediction* concerning the values of  $\pi(x)$  within Vega's and Chernac's tables. It is also a prediction concerning every future PNT.

(b) The way followed by Legendre in proposing his own hypothesis (PNH-2) is more uncertain. However, J.W.L. Glaisher and J. Glaisher carefully studied this question, concluding that:

"It would thus appear that Legendre, having led by analytical considerations just mentioned or otherwise, to a formula of the form  $x/A \log x - B$ , determined the values of A and B empirically by means of the enumerations: the value of A would be at once found to be unity".<sup>55</sup>

As a matter of fact, Legendre published an *Addenda* to his *Théorie des nombres* in 1816, in which he extended his calculations up to 1.000.000 and confirmed the validity of PNH-2. Therefore, we can conclude that as much Legendre as Gauss employed empirical methods and frequently used PNT in the aim of stating, modifying and corroborating their hypotheses.

(c) Wenn Gauss knew Legendre's research on PNP, he contrasted both analytical expressions with the available tables in 1849. Gauss verified the value of Legendre's constant for higher intervals of numbers than Legendre did it and observed a very important fact: the value of Legendre's constant decreases when more extensive tables are employed. The commentary of Gauss about this empirical fact, entirely new at this time, can be considered to be the origin of a new basic concept in Analytical Number Theory: the asymptotic convergence,  $O(x)$ , which was formally defined by Landau many years later. Littlewood proved that  $\pi(x)$  is an oscillatory function and defined the new concept of asymptotic convergence of functions in NT.

(d) When PNH-1 and PNH-2 became a *mathematical conjecture*, which were received by the community of mathematicians as a consistent problem, new hypotheses and methods proliferated. Encke, Chebychev, Riemann, Meissel and many other mathematicians contributed with their proposals, trying to prove some rigorous statement about PNP. At this time, two kinds of general methods were usually employed: the deductive one's and, simultaneously, the improved empirical methods. Some mathematicians employed more powerful analytical methods. The rest of mathematicians continued to support only elementary methods: a strong controversy began between number theorists. However, both factions accepted the final verdict of prime number tables:

---

<sup>55</sup> J. Glaisher 1883, p. 67.



formulas had to agree with empirical facts. Therefore, it is not an excessively audacious claim to consider this kind of research as an experimental one, from a methodological point of view.

(e) What could be the epistemological function of PNT in developing this kind of research? Undoubtedly, prime number tables were at the origin of Gauss's and Legendre's hypotheses. PNT prompted also the emergence of several mathematical concepts, such as  $\text{Li}(x)$  and  $O(x)$ . Research on PNP was largely promoted by the advances accomplished by constructing PNT. Each proposed hypothesis had to be verified by a comparison to PNT. Consequently, we can conclude that prime number tables were employed by number theorists just as experimental scientists used their empirical measures or their symbolic expressions of facts in physics, chemistry or social sciences.

(f) Finally, it is important to remark that I do not argue in favour of the existence of an empirical basis in mathematics, nor even in number theory. This kind of ontological problems exceed my present purpose, which is strictly limited to some epistemological, historical and methodological matters. My only claim is to the importance of prime number tables in NT. Using tables counting prime numbers and tables measuring empirical data in order to prove or to disprove hypotheses should be considered as the same method with different applications.

## Historical Aspects of the Foundations of Error Theory

EBERHARD KNOBLOCH (Berlin)

"The method of least squares is the automobile of modern statistical analysis: despite its limitations, occasional accidents, and incidental pollutions, it and its numerous variations, extensions and related convergences carry the bulk of statistical analogies and are known and valued by nearly all" (Stigler 1981, 465).

My intention is not to repeat the derivation of the technical details. This has been done by many authors during the last decades, very recently by Jean-Luc Chabert (1989). My intention is to concentrate on the question: Why was there such a long dispute over the foundations of this method? Why were there so many proofs of the method during the whole 19th century, with such prominent mathematicians as Laplace and Gauss involved? This is indeed a fascinating story concerning mathematical and physical concepts and methods like proofs, rigour, strictness, experience, observations. I would like to discuss the following four problems:

1. Systematical and methodological aspects.
2. The most important proofs and the contemporary criticisms of them.
3. The hypothesis of elementary errors.
4. The principle of the arithmetical mean.

### *1. Systematical and Methodological Aspects*

In 1853 the French financial inspector Irénée Jules Bienaymé said in his article "Considerations which will confirm Laplace's discovery of the probability law in the method of least squares" (1853,310): "The probability calculus is the first step of mathematics beyond the region of absolute truth". And indeed about 500 essays of the whole 19th century dealt with the foundation of error theory, especially with the method of least squares. They show how difficult it was to find such a foundation. Some authors like James Ivory tried to give it without using probability theory but – at least according to his critics – he did not succeed.

The discussion continued through the whole century. Each of them, Pierre Simon Marquis de Laplace, and Carl Friedrich Gauss, have claimed to give two proofs of the method of least squares. Neither Gauss nor Laplace, nor any other author, convinced all the critics who examined the methods. Several fundamental questions remained unsettled which all originated from the fact that the interrelations between physical reality and mathematical description were unavoidably concerned:

1. What is the "true foundation of the method?" (Ellis 1844, 204)
2. Which are the admissible fundamental principles and assumptions?
3. What can be classified as a proof of the method or what is a rigorous demonstration? What is a mere consideration?
4. What did an author prove, suppose he proved something?
5. Which mathematical description of reality is appropriate?

Let me explain these four problems.

### *1.1. What Is the True Foundation of the Method?*

This question concerns the metaphysics which underlies the method of least squares. Gauss himself used the expression "metaphysics" in his correspondence with the astronomer Friedrich Wilhelm Bessel dating from 1839 (Gauss 1880, 523). Four years later the secondary school professor Karl Gustav Reuschle took up this expression and distinguished between two concepts (1843, 333):

- (i) The metaphysics, that is, agreement upon and analytical representation of the fundamental concepts, and
- (ii) The algorithm, that is, the development of the system of formulas and of calculation devices.

The professor of political economy Francis Ysidro Edgeworth who was heavily concerned with error theory, said in (1884, 210):

"How much such a foundation will support, to what height it is expedient to carry an arithmetical calculation founded thereon, is a question to be determined by that unwritten philosophy and undefinable good sense which, in the order of scientific method, precedes the application of calculus and is prior to a priori probabilities".

The problem consisted in how to apply this philosophy or undefinable good sense. There was no clear answer. Even in 1892 Paolo Pizzetti (1892, 15) remarked:

"The false applications of error theory and its misunderstood results are

mostly responsible for the uncertainty which still exists with regard to the philosophical foundations of this theory".

No wonder that Morgan William Crofton designated Laplace, Gauss, and Poisson, "who may be called the founders of the Theory of Errors", as "great philosophers" (1870, 176). Thus, whenever authors tried to evaluate the demonstrations presented up to their time, such as Johann Franz Encke in [1832], Robert Leslie Ellis in [1844], James Whitbread Lee Glaisher in [1872], or Pizzetti in [1892], they came to very different conclusions because they did not agree with one another on the underlying philosophy. This is especially true with James Ivory. He wanted to lay before the reader that particular view concerning the principles of this method which appears most natural and philosophical (1825, 87). In his opinion all proofs by means of the doctrine of probabilities were "entirely supposititious and mathematical" and therefore insufficient and unsatisfactory (1826, 165). In order to avoid all these "precarious suppositions" we have to set out by demonstrating the most advantageous solution from the nature of the equations of condition. As a consequence the whole theory will follow "naturally" and will be placed "on its proper foundation" (1825, 88).

### *1.2. Which Are the Admissible Fundamental Principles and Assumptions?*

Already Gauss and Reuschle explicitly claimed the right to use certain hypotheses, certain assumptions (Gauss 1809a, 103; Reuschle 1843, 356). But this right remained controversial. This applies especially to the so-called hypothesis of elementary errors and to the principle of the arithmetical mean which we shall discuss more precisely a little bit later. Donkin said in 1851 that the probability of every hypothesis depends upon the state of information presupposed concerning it. If the actual law of facility of errors is not known, every solution must involve an assumption. This assumption should have at least three characteristics:

1. to be the most simple,
2. to be the least arbitrary, and
3. to be the most in accordance with common notions and experience.

But all three criteria remained controversial, especially the simplicity and arbitrariness of an assumption. Sometimes the authors, for example Donkin, until lately, did not clearly apprehend which was the assumption really involved in their proof. Donkin's assumption was: The knowledge gained from a number of observations is the same in kind as that gained from a single observation. He was inclined to think this assumption, in itself, was more simple and natural than any other, but he confessed that this is a matter of opinion (1851, 59).

This statement should be kept in mind. Just because many authors admitted that the acceptance or rejection did not depend on objective criteria, but on very subjective decisions, on the opinion (Donkin), on the feeling, intuitive understanding (*Gefühl*) of the author (Encke 1844, 222), on the undefinable good sense (Edgeworth 1884, 210), they could not come to any agreement. There remained "fissures" to be filled up by legitimate conjecture, as Edgeworth said. But the legitimacy of a special conjecture remained questionable. What is more, even the legitimacy of a special conjecture at all was contested by Ivory (1826, 161). He was convinced that the "real principles alone concerned" had to be free from "everything conjectural or tentative".

### *1.3. What Can Be Classified as a Proof of the Method?*

After these explanations we understand why different authors classified one and the same justification of the method of least squares in extremely different manners. This applies for example to Gauss' second proof. Ellis (1850, 321) called it "perfectly rigorous", while Mansfield Merriman said in 1877 that the proof is entirely untenable and, he believed, that it was only followed by the German geodesist Friedrich Robert Helmert in 1872 (Merriman 1877b). Merriman did not accept that the proof took for granted that the mean value of the sum of the squares of the errors may be used as a measure of the precision of the observations.

This applies, too, to James Ivory's proof of 1826 [Ivory 1826]. Robert Leslie Ellis called it (and its predecessors) "not at all satisfactory" (Ellis 1844, 204). Merriman "still more absurd than those of 1825" (Merriman 1877a, 177) or "the most unsatisfactory of all" (Merriman 1877b), while in 1830 and still in 1840 Louis Benjamin Francoeur called it a "perfectly valid demonstration" (Merriman 1877a, 178). His first demonstration (1825) was classified by Ellis (1844, 217) as a "vague analogy", because Ivory, who was mainly interested in the applications of mathematics to physical problems, tried to establish an analogy between the influence of the error  $e$  on the value of the correction  $x$  with a lever in mechanics which is to produce a given effect. While Ivory rejected the probability theory as a mathematical means of proving the method, because such proofs are founded on arbitrary suppositions and so are inconclusive, Glaisher on the contrary, was aware that any reasoning without recourse to this theory had to be inconclusive (Glaisher 1872, 83).

### *1.4. What Did the Author Prove, Supposing He Proved Something?*

There was no agreement on that what the authors proved. Donkin especially

underlined two aspects: (i) Gauss proved the method to be a "very good" method, though he did not prove that it is the "best" method, but he had not shown that it gives the most probable result. He rigorously demonstrated what he professed to demonstrate, but he did not claim "to demonstrate the method of least squares, in the sense in which these words would be commonly understood without explanation" (1851, 58). (ii) To prove that "a required probability is to be calculated as if a certain hypothesis were known to be true is a perfectly different thing from proving that that hypothesis is true, or from proving anything about the probability of its truth at all" (Donkin 1851, 57).

### *1.5. Which Mathematical Description of Reality Is Appropriate?*

All four foregoing questions depend more or less on the mathematical description of reality, because the method of least squares is a veritable pivot between pure and applied sciences (Chabert 1989, 26). There are two aspects which play the main role in the discussion between the authors dealing with this method:

- (a) the methodological aspect that we have to determine whether something is a psychological postulate or a result of experience (de Morgan 1864, 409).
- (b) the theoretical aspect that the mathematical theory should express, at least approximately, what generally does occur "in rerum natura" (Crofton 1870, 176).

The first aspect implies two kinds of truth: (i) the a priori truth of a priori mathematical assumptions (Ellis 1844, 205), for example the principle of the arithmetic mean; (ii) the practical truth (Crofton 1870, 177) which is a question of facts. It depends on an inquiry, not into what might be, but what is. For example, the error law is so far practically true as far as the underlying hypothesis (for example of elementary errors) is in accordance with fact. The confidence in the permanence of nature implies a conviction that the effect of fortuitous causes will disappear on a long series of trials (Ellis 1844, 205). This conviction leads to the principle of the arithmetic mean.

The second aspect helps to avoid the mistake that mathematical fictions are intermingled with reality. The application itself of the probability theory to the study of errors of observation are based on a fiction which ought not be made reality, as Bertrand pointed out (1889, 215).

But in order to derive conclusions which correspond to, and represent, outward reality we have to know something. "Mere ignorance is no ground for any inference whatever" as Ellis said (1850, 325) when he rejected Herschel's (1850) demonstration. For Herschel justified the assumption that the law of

errors is in all cases the same, by our ignorance of the causes on which errors of observation depend.

## *2. The Most Important Proofs and the Contemporary Criticism of Them*

We can enumerate at least eighteen proofs. Thirteen of them are enumerated by Merriman (1877a, 153f; 1877b):

0. Adrien-Marie Legendre (1805) (publication of the method without proof)
- 1./2. Robert Adrain (1808)
3. Carl Friedrich Gauss (1809)
4. Pierre Simon Marquis de Laplace (1810)
5. Pierre Simon Marquis de Laplace (1812)
6. Carl Friedrich Gauss (1823)
- 7.-9. James Ivory (1825)
10. James Ivory (1826)
11. Gotthilf Hagen (1837)
12. Friedrich Wilhelm Bessel (1838)
13. William Fishburn Donkin (1844)
14. John Herschel (1850)
15. William Fishburn Donkin (1857)
16. Peter Guthrie Tait (1865)
17. Morgan William Crofton (1870)
18. Paolo Pizzetti (1892)

I shall concentrate on Gauss' and Laplace's proofs, but I shall say a few words on some others, too, without entering into priority questions (see Stigler 1981).

### *2.1. Legendre's Publication of the Method (0.)*

Legendre (1805) did not give any proof (Schneider 1981, 144, 151). He underlined the advantages of his method. There is no principle, he said, which can be proposed for this subject which is more general, more precise or more simply applicable. The principle of the arithmetical mean is but a simple consequence of his general method.

### *2.2. Gauss' First Proof of the Method (3.)*

When Gauss wrote to H. C. Schumacher in 1844, he judged the first kind of foundation to be this procedure: to base the method solely on the "principles of

practicality" (Gauss 1862, 371). Apart from Adrain's unsatisfying article (1808) which was probably published in 1809, Gauss gave the first, that is probability-theoretical proof of the method in (1809a). He reversed in a certain way Legendre's statement saying:

"If the generally accepted principle of the arithmetical mean is necessarily true, there are two consequences:

1. A certain probability law is necessary.
2. The only method of combining the equations obtained by observations is the method which minimizes the sum of the squares of these expressions".

He determined the probability law by means of the maximum likelihood principle and got the normal distribution, that, as he proved, demonstrates that among unimodal, symmetric and differentiable distributions  $\phi(x-x_0)$  there is a unique distribution (the normal one) for which the maximum likelihood estimator  $\hat{x}$  of the location parameter  $x_0$  coincides with the arithmetic mean.

If the law of distribution of errors  $\Delta$  is given by

$$\phi(\Delta) = \frac{h}{\sqrt{\pi}} e^{-h^2 \Delta^2}$$

then the function

$$\Omega = h^\mu \pi^{-\frac{\mu}{2}} e^{-h^2(v_1^2 + v_2^2 + \dots + v_\mu^2)}$$

attains its maximum value if

$$v_1^2 + v_2^2 + \dots + v_\mu^2 = \min.$$

where  $\mu$  is the number of observations,  $v_i$  are the differences between the observed and calculated values of given linear forms of the unknowns sought (Sheynin 1979, 31).

Gauss claimed this foundation of the method as completely his own (1809b, 205). But he admitted in the self-announcement of the "*Theoria combinationis observationum erroribus minimis obnoxiae*" (1821, 194) that it was still unsatisfactory in a way:

- (i) The foundation depends exclusively on the hypothetical form of the probability law of errors. If we abandon this form, the values obtained by the method of least squares are no longer the most probable values. The same would be true with the arithmetical mean in the simplest of all cases. As the law of the probabilities of errors of observation remains always hypothetical, he applied this theory to the most plausible law.
- (ii) He mentioned in a private letter to Bessel dating from 1839 another aspect which in his opinion impaired his own proof (Gauss 1880, 523f):



"It is less important to calculate the value of an unknown quantity, the probability of which is maximal (it is by all means only infinitely small), than to calculate the value which enables us to play the least disadvantageous game".

He took the idea from Laplace to give a game-theoretic foundation of error theory. His two own objections to his first foundation did not coincide with the many objections made by other contemporary and later authors:

1. Principal objection (to all other demonstrations, too, of the formula  $\phi(x) = Ce^{kx^2}$ ): The basis of all demonstrations are those postulates that have to be satisfied by the function  $\phi$  (Laplace 1816, 500; Bienaymé 1852, 36f; Pizzetti 1892, 216).
2. There is no reason for supposing that, because the arithmetical mean would give the true result if the number of observations were increased without limit, it must give the most probable result when the number of observations are finite (Ellis 1844, 207). In other words, Gauss' proof does not give the most probable result.
3. The most advantageous value is not the most probable value (Peirce 1870; Edgeworth 1883; Bertrand 1889; Pizzetti 1892, 215).
4. We have no right to assume, as an axiom, that the arithmetic mean is the most probable result of every series of direct observations (presumed to be equally good) of the same quantity (Glaisher 1872, 84; Merriman 1877a, 165).
5. "It is not recognized that the probability of a definite error  $x$ , is an infinitesimal" (avoided by some later writers) (Merriman 1877a, 165).
6. The distinction between true errors and residuals (or calculated errors) is not sharply drawn:  $\phi$  is not strictly a "law of facility of errors" but of distribution of residuals (Merriman 1877a, 165).
7. (*difetto fondamentale*) We have no complete liberty when choosing the observation values  $l_1, l_2, \dots, l_n$ . This is a purely mathematical fiction (Pizzetti 1892, 213).
8. The proof of the formula  $\phi(x) = Ce^{kx^2}$  is rendered completely invalid if the arithmetical mean represents only approximately the most probable value as Gauss indicates (Gauss 1809a, 101; Pizzetti 1892, 214).
9. It is an absolutely arbitrary assumption that the relative probability of an error can be expressed by a continuous function of this error (Pizzetti 1892, 215).

These chronologically ordered objections can be classified into three groups:

- (a) different aspects of the principle of the arithmetical mean (objections 2,4,8);

- (b) existence and properties of the error function (objections 1, 6, 9); and
- (c) the mathematical description of experimental facts (3, 5, 7).

Already Gauss himself had commented upon the hypothetical form of the probability law of errors. Donkin agreed with him insofar by saying (1857, 160):

"The utmost which any such process can pretend to establish is, not that the unknown law of facility of error *is* expressed by a function of this form (which would be manifestly an absurd pretension), but that *the law being unknown* the most probable result is to be obtained by proceeding *as if* it were known to have the form in question".

A main problem was, and remained, the use of the principle of the arithmetical mean. Bienaymé being an adherent of Laplace, did not accept this arbitrary hypothesis as being the a priori hypothesis of the necessity of the minimum of the squares (1852, 36). For him the proof was restricted to those special cases where the assumed probability law of errors is to be found in the observations.

Laplace had in view his own proof when he criticized Gauss' proof:

- (i) Gauss had not shown that this principle provides the most advantageous result. But that, to be sure, had not been Gauss' intention.
- (ii) He criticized (as Gauss himself had done) the dependence of the foundation on the probability law of errors of observations.

### 2.3. Laplace's First Proof of the Method (4)

Laplace's own aim was to demonstrate that the method of least squares is the most advantageous method and that it is a priori (Laplace 1812, 342). He underlined the fact that his proof was independent from any special law of probability (Laplace 1816, 500).

The second statement was crucial for Glaisher (1872,92): whatever strictly philosophical basis the subject has, must be therefore attributed to Laplace according to Glaisher.

To be sure, this statement was only valid provided that there were very many (infinitely many) observations. Indeed, the independence from the law of errors resulted from the theorem known today as central limit theorem, which was rigorously proved by Chebychev in 1887 (Chabert 1989, 16). He mentioned at the same time that his analysis was based on the hypothesis that there are equal facilities (relative probabilities) of positive and negative errors, though he considered the case, too, that these facilities were different.

What happened to Laplace's first proof?

1. Gauss criticized that this proof did not explain what we have to do in the usual case of a moderate number of observations (Gauss 1821/22, 194). The practitioner and astronomer Encke agreed with Gauss (Encke 1832, 74) and so did many others.
2. Moreover, Gauss criticized Laplace's definition of the mean maximal error for which we have to be prepared: Laplace had defined it by the sum of the products of each error multiplied by its probability (Laplace 1812, 324). This mean is not proportional to the limits of the errors. Therefore the result was – in Gauss' opinion – not rigorous.  
Gauss underlined in his second proof (1823, 6, 20) the advantage of his definition over that of Laplace (it is the modern notion of variance):

$$m^2 = \int_{x=-\infty}^{x=+\infty} x^2 \phi(x) dx$$

3. Many authors were irritated at the obscurity of Laplace's proof. John Herschel called Laplace's analysis "exceedingly complicated" (Herschel 1850, 11), Joseph M. de Tilly found Gauss' theory "infiniment plus simple" than Laplace's proof (de Tilly 1874, 138). Finally Ellis said in his (1844, 212):

"It must be admitted that there are few mathematical investigations less inviting than the fourth chapter of the *Théorie des probabilités* which is that in which the method of least squares is proved".

Other authors hinted at hidden assumptions used by Laplace:

4. Laplace supposed the law of error the same at each observation (Ellis 1844, 212).
5. Laplace took the first step in his method upon the following principle: Where the form of a function is completely unknown, it is allowable to assume that form which is most convenient for the purpose of calculation (Edgeworth 1884, 209).
6. One of Laplace's principles was the assumption that an error of observation is produced by "the algebraic combination" of many independent sources of errors (Airy 1861, 7; Tait 1865, 139). Airy added, that this was not the language of Laplace.

Bienaymé (1852) defended Laplace's proof as well as he could: It is simply false to apply the method of least squares to only a few observations. He rejected Gauss' criticism of Laplace's definition of the mean maximal error: Gauss' criterion of precision (the mean of the square of the errors) is not proportional to the limits of the errors, either. He rejected the criticism of the too great complexity: those who have looked for the spirit of this analysis have wasted their efforts to replace the calculations of Laplace by that which is called

"simple demonstrations" or "popular proofs".

Nevertheless, Ellis was able to diminish greatly the mathematical difficulty by substituting Fourier's theorem for Laplace's "theory of combinations" and use of imaginary symbols (Ellis 1844, 212).

#### *2.4. Laplace's Second Proof of the Method (5.)*

Laplace's second proof differed from the first by an important peculiarity. Now he considered the errors of observations as losing in a game in which we cannot win because we can never obtain more than the truth. He identified the losses with the absolute value of the error (Laplace 1812, 324).

Encke criticized that the estimation of the losses according to the magnitude of the error implied an unproven assumption concerning the interrelation between the errors (Encke 1832, 74).

#### *2.5. Gauss' Second Proof of the Method (6.)*

Gauss took up the game-theoretic approach in his second proof dating from 1821–1823. It is no wonder that Encke adhered to his first proof, because he rejected this kind of reasoning already in Laplace's second proof. Nevertheless, Gauss criticized Laplace for having chosen the absolute value of the error as a measure of the losses. This was too arbitrary in his opinion. His criterion was simplicity: among all infinitely many functions which are always positive and can express the magnitude of the losses by a function of the error, the simplest function has to be chosen. This function is incontestably the square.

Therefore, Gauss chose the variance as a measure of precision and adjusted the observation according to the principle of least variance. Later on he accepted only this foundation, though he admitted that the choice of the squares was purely arbitrary (Gauss 1880, 524): "Without the known great advantages of this choice we could choose any other function which satisfy those conditions".

What happened to Gauss' second proof?

Ellis accepted it without any restriction. In his opinion it was a perfectly rigorous demonstration (1850, 321): "Nothing can be simpler or more satisfactory than this demonstration", he said in (1844, 216):

"The proof is as simple as possible, free from any analytical difficulties, and independent from the number of observations".

But there was no lack of criticism:

Encke criticized the arbitrariness as well as the principle of simplicity (Encke 1832, 74).

Bienaymé denied that this was a proof. He believed it to be only considerations ("considérations") (1852, 37). He denied the arbitrariness of the choice of the squares. For a great number of observations the most advantageous method leads necessarily to the method of least squares.

Later authors like Glaisher (1872) or Merriman (1877a) criticized Gauss' measure of precision. As a consequence, the proof rested upon an arbitrary assumption. Merriman (1877a, 174) even said: "It is but little more than a beginning of the question to assume that the mean of the squares of the errors is a measure of precision". Thus in his opinion the proof was completely untenable.

### *3. The Hypothesis of Elementary Errors*

There is a close relation between the derivation of the error function or density function

$$\phi(x) = Ce^{kx^2} = \frac{h}{\sqrt{\pi}} e^{-h^2 x^2}$$

and the so-called hypothesis of elementary errors. Pizzetti called it the fundamental hypothesis, on which he had based his own considerations (1891, 224). But only after several decades was this hypothesis developed in full generality and widely accepted after the assumptions had been extended step by step.

The first author who looked for the error law assuming explicitly that every error is produced by a great number of independent causes was Thomas Young in 1819. He wrote to Henry Kater, saying (1819, 71):

EY "The combination of a multitude of independent sources of error, each liable to incessant fluctuation, has a natural tendency, derived from their multiplicity and independence, to diminish the aggregate variation of their joint effect".

Bessel's pupil Heinrich Ludwig Hagen expressed this principle more precisely in his (1837, 28).

EH "The observational error is the algebraic sum of infinitely many elementary errors which all have the same value and which can be equally easily positive or negative".

Hagen believed that this assumption is immediately explained by an analysis of the measuring method and the combination of the device. The number of elementary errors is the more increasing, the more sources of errors are taken into account. But EH was apparently rather restrictive and was rejected as a "questionable hypothesis" by Joseph Dienger in his (1852).

One year later Bessel published his "Inquiries into the probability of observational errors". He wanted to base his analysis on the mode of the origin of the observational errors which depends on their sources. Therefore he analyzed the probability of errors assuming the following hypothesis:

EB An error is produced by the combination of several sources being independent from one another. Every source has such an effect that positive and negative errors of the same magnitude have the same probability.

Thus Bessel assumed errors of different magnitudes. He was a sufficiently good practitioner to know that every trial must be necessarily impotent to recognize, in general, the law which underlies the method of least squares as that which occurs in reality. There are conditions which are mathematically possible and can be practically fulfilled, but that imply completely different laws.

Morgan William Crofton believed even this hypothesis to be still too restrictive. In 1870 he published his article "On the proof of the law of errors of observations". He presupposed that positive and negative values of each error are not assumed equally possible, and that each minute simple error followed its own unknown law, expressed by different unknown functions of the utmost generality.

He said (1870, 177):

"As far as this hypothesis is in accordance with fact, so far is the law practically true. Fully to decide how far this hypothesis does agree with facts is an extremely subtle question in philosophy, which would embrace not only an extended inquiry into the laws of the material universe, but an examination of the senses and faculties of man, which form an important element in the generation of error".

Without pretending to enter on a demonstration of the truth of this hypothesis, he wanted at least to convince the reader of its reasonableness in certain large classes of errors of observations.

All methods applied up to then were deficient in generality as he said. He wanted to give a proof which was as general as possible by excluding the term probability and considering solely the frequency or density of the error viewed as a function of its magnitude.

Without having any antecedent knowledge of the peculiar property of combination of the errors, he derived the following function of error of a system of minute, combined errors:

$$Y = \frac{N}{\sqrt{2\pi(h-i)}} e^{-\frac{(x-m)^2}{2(h-i)}}$$

$m = \alpha + \beta + \gamma + \dots$  = sum of mean errors

$h = \lambda + \mu + \nu + \dots$  = sum of mean squares of errors

$i = \alpha^2 + \beta^2 + \gamma^2 + \dots$  = sum of squares of mean errors

$N$  = number of observations taken, affected with the compound error.

He only adopted the "usual axiom" as he said:

*Axiom:* No function can represent a finite error unless the mean value, mean square, mean cube etc. of the finite error  $y = F(x)$  are finite.

The probability of an error being found to lie between  $x$  and  $x + dx$  is

$$\frac{1}{\sqrt{2\pi(h-i)}} e^{-\frac{(x-m)^2}{2(h-i)}} dx$$

If positive and negative errors in the observations are equally probable, as generally can be secured in practice, at least approximately, then  $m = 0$ ; that is the sum of the mean values of the elementary component errors vanishes, and the probability is expressed by the usual value:

$$\frac{1}{c\sqrt{x}} e^{-\frac{x^2}{c^2}} dx$$

Thus Glaisher summarized the examinations of the proofs by saying (1872, 120):

"It seems to me that the only sound philosophical basis on which the law of facility  $e^{-h^2x^2}$  rests, is the supposition that an actual error is formed by the accumulation of a great number of small errors due to different independent sources, and subject to the arbitrary laws of facility  $\phi_1(x)$ ,  $\phi_2(x)$  ...".

Nevertheless this understanding was not at once accepted, especially because the development was not generally known which took place after Hagen's publication. In 1871, one year before Glaisher published his article, Lorenz Lindelöf wrote his review of V. Neovius' textbook on the method of least squares. Lindelöf criticized the author for having adopted Hagen's hypothesis. Lindelöf rejected it for two reasons:

1. It is not justified by any, even plausible, consideration.
2. It hasn't even got the formal advantage that the probability law of errors can be deduced from it without an additional hypothesis.

He wanted to show that one can derive the law

$$\frac{\phi(v)}{\phi(0)} = e^{-h^2v^2}$$

by means of Hagen's hypothesis only by assuming an arbitrary presupposition. Otherwise Hagen's opinion leads nowhere. This presupposition consists in the assumption that there is such a relation between the maximal error  $V$

and the elementary error  $\alpha$  that the product  $V\alpha$  converges to a limit which is unequal to 0 (Knobloch 1985, 573–575).

Three years later Joseph Marie de Tilly, an opponent of Hagen's hypothesis, spoke of an important objection to the method (de Tilly 1874) with regard to Lindelöf's argument.

#### 4. *The Principle of Arithmetical Mean*

When in 1809 Gauss gave a probability-theoretical foundation of the method of least squares, he said (1809,101):<sup>1</sup>

"The following hypothesis is generally accepted as an axiom: If any arbitrary quantity is determined by several direct observations made under the same circumstances and equally diligently, the arithmetical mean provides the most probable value although not absolutely rigorously, yet very nearly so that it is always the safest method to adhere to it".

And a little later:<sup>2</sup>

"This principle (that is the method of least squares) has to pass as an axiom with the same good reasons with which the arithmetical mean of several observed values of the same quantity is chosen as the most probable value".

What had Gauss said? The opinions were divided. Ellis (1850), Herschel (1850), Glaisher (1872) believed that Gauss had assumed that the arithmetical mean is the most probable value. Herschel required a proof of this assumption, therefore he did not accept the following explanations as a proof. On the other side, William Chauvenet said in 1868 that this principle is the most simple and obvious, and might well be received as axiomatic (1868, 475).

Glaisher underlined that we have no right to assume the principle of the arithmetical mean as an axiom and reproached Gauss for not having tried to prove this principle. Though experience has shown that the arithmetical mean provides very good results it cannot be shown that it provides the best possible (1872, 84).

But Glaisher qualified his statement. Gauss' view was only that the arith-

---

1 Axiomatis scilicet loco haberi solet hypothesis si quae quantitas per plures observationes immediatas, sub aequalibus circumstantiis aequalique cura institutas, determinata fuerit, medium arithmeticum inter omnes valores observatos exhibere valorem maxime probabilem, si non absoluto rigore, tamen proxime saltem, ita ut semper tutissimum sit illi inhaerere.

2 Hocce principium, quod in omnibus applicationibus mathesis ad philosophicam naturalem usum frequentissimum offert, ubique axiomatis eodem iure valere debet, quo medium arithmeticum inter plures valores observatos eiusdem quantitatis tamquam valor maxime probabilis adoptatur.



metic mean is practically the best mode of combining simple observations and that experience has justified its adoption by the accuracy of the results obtained. Gauss was very far from asserting, as a result deduced from the theory of probability, that the arithmetic mean is the most probable value of the quantity observed. Already, two years earlier M. W. Crofton took this intermediary view when stating (1870, 176): The principle of the arithmetical mean "is not an axiom, but only a convenient rule which is generally near the truth". According to Crofton, Gauss himself was very far from asserting the assumption of being an axiom (Glaisher repeated these words, nearly word by word, without mentioning Crofton). He did not give his proof as anything more than hypothetical.

The principle of the arithmetical mean was the basis of Gauss' proof published in 1809. This did not imply its uncritical application. Already in 1805 Bessel wrote to Gauss: "By all means the mean must not be taken blindly without a foregoing examination" (Gauss 1880, 26). In 1832, the astronomer J. F. Encke tried to improve Gauss' proof, saying that all theoretical foundations of the method of least squares presented up to then had not achieved their purpose (Encke 1832). He did not accept the principle of the arithmetical mean as an unproven axiom. Thus he saw two alternatives:

1. The principle is proven by means of simpler axioms which are not to be demonstrated further.
2. Centuries-long experience takes the place of the rigour for the theoretical reasoning, which is lacking.

Encke chose the first alternative. He wanted to demonstrate that the principle is the most probable, or at least, the only completely consistent method which has to be chosen preferably. Gauss chose the second alternative. But what about the general validity of this experience? We shall discuss this problem below.

#### *4.1. The First Alternative or the Reduction to Simpler Axioms.*

According to his own words, Encke took as a basis the following two hypotheses:

- H1 The probability of an error depends only on its magnitude. It does not depend on its sign, that is, the most probable or most advantageous value must be an even function of the observations.
- H2 We obtain the most probable result if we combine the single observations in groups according to the right principles and take together only the results of the combinations without taking further into account the single observations.

Pizzetti (1892, 196) when analyzing Encke's proof pointed out a third assumption:

A1 In the case of two observations the arithmetical mean is the most advantageous value.

To be sure, Encke deduced this assumption from another assumption:

A2 All observations are completely uniform (equally reliable).

We have to draw attention to a further assumption which Encke mentioned as a matter of course, not as an explicit assumption (Encke 1832, 75):

A3 The error law is unknown.

Encke mentioned H2 only in his reply to Reuschle's criticism (Reuschle 1843). His proof ran as follows:

Let  $a, b$  be two observations,  $x$  the quantity which is directly to be determined, or its assumed value,  $x-a, x-b$  the errors. Then  $x = \frac{1}{2}(a+b)$  according to H1, A2.

If we have three observations  $a, b, c$ , we get

$$(1) \quad x = f(a, b, c), \quad f \text{ symmetric} \quad (A2)$$

$$(2) \quad x = \psi\left(\frac{1}{2}(a+b), c\right) = \psi\left(\frac{1}{2}(a+c), b\right) = \psi\left(\frac{1}{2}(b+c), a\right) \quad (H2)$$

We put

$$(3) \quad s = a + b + c \quad \text{Then we obtain}$$

$$(4) \quad x = \psi\left(\left(\frac{1}{2}s - \frac{1}{2}c\right), c\right) = \psi\left(\left(\frac{1}{2}s - \frac{1}{2}b\right), b\right) = \psi\left(\left(\frac{1}{2}s - \frac{1}{2}a\right), a\right)$$

$$(5) \quad x = \psi(s, c) = \psi(s, b) = \psi(s, a)$$

$c, b, a$  have to disappear because of (1), if we develop  $\psi$ . Thus

$$(6) \quad x = \psi(s)$$

In the special case

$$(7) \quad a = b = c$$

we get  $x = a$  because there is no possible choice.

Thus we get from  $a = \psi(3a)$

$$(8) \quad \psi = \frac{1}{3} \quad \text{or} \quad x = \frac{a+b+c}{3}$$

The proof can be completed by complete induction. What happened to Encke's proof? Nearly every step implied a difficulty that is provoked an objection

though there were authors like William Chauvenet in 1868 who accepted it without any modification. He adhered even to the same function sign with regard to the transition from (4) to (5). But Reuschle was right in pointing to the fact that these two equations imply two different modes of functional relations.

Reuschle was the first of a long series of authors who criticized Encke's proof for different reasons. Already the aim of the proof was criticized by J. Bertrand (1889, 171), because it mixed up two things which are independent from each other:

1. In the presence of several measuring results of an only quantity, it is the best decision to get the mean.
2. The mean of several measures is the most probable value.

Other authors objected to special assumptions used in the proof.

*Objection 1* (Reuschle 1843): Why should one introduce the quantity  $s$  by equation (3)? We could introduce any other symmetric and homogeneous function of  $a$ ,  $b$ ,  $c$  for example

$$s = a^m + b^m + c^m,$$

eliminate  $a$  or  $b$ , or  $c$  from the function  $\psi$ , for example we could substitute

$$c = \sqrt[m]{s - a^m - b^m}$$

and add the remaining conclusions.

Encke defended his method by saying that such a substitution is impossible, supposing we apply the hypothesis H2. In this case the necessary knowledge of the single quantities  $a$ ,  $b$  for such a substitution is not presupposed (Encke 1844).

*Correction 1* (Reuschle 1843)

Encke did not comment on Reuschle's remark that we have to assume a characteristic of the function  $\psi$  instead of speaking of lack of choice:

H3 The even function looked for by H1 is reduced to one value if all observational values are equal to this value.

Reuschle believed that one could deduce by means of H1, H3, that the arithmetical mean is the simplest and therefore the most natural, as well as the most appropriate, which is the most plausible mean among all suitable functions.

Thus, Donkin stated in (1851, 57): On the understanding that H3 holds it has neither been proved that the mean is the most probable result, relative to

this state of information, nor has it clearly been proved that it is not. The question is perfectly open.

Therefore it seemed to be reasonable to look for other proofs.

*Objection 2 (Schiaparelli 1868)*

In 1868 Schiaparelli again took up the problem. He objected as did Glaisher (1872) and Tilly (1874) later on, to the evidence of hypothesis H2: Why should the result  $x$  (see above equation (2)) be a function of the half sum  $\frac{1}{2}(a+b)$ , only because this half sum is the most plausible value suppose the third observation is lacking?

*Objection 3 (Schiaparelli 1868)*

Nothing indicates that the correction which has to be applied to the result  $\frac{1}{2}(a+b)$ , considering the third observation  $c$  does not have to be a function of the preceding results  $a, b$ .

Though he presented two proofs by means of different hypotheses, he did not dare to maintain that he had been successful in doing so: For a great variety of judgements is possible in such a delicate matter.

*Correction 2 (Schiaparelli 1868)*

A2, H1 and the following two "evident hypotheses":

H4 If we add a constant quantity to all observed values the result has to increase by this constant.

H5 If we multiply all observed values by  $k$ , the result has to be multiplied by  $k$ , too.

*Correction 3 (Schiaparelli 1868)*

A1

H6 If we consider the quantities  $a, b, c, d$  as equal with regard to the result, all their similar combinations like  $\frac{1}{2}(a+b), \frac{1}{2}(a+c)$  etc. are to be considered as equal and of equal precision.

Glaisher's objection dating from 1872 was similar to objection 2.

*Objection 4 (Glaisher 1872)*

Why should the most probable result from  $a, b, c$  be a function of the most probable result from  $a$  and  $b$ , and from  $c$ ?

One year after Schiaparelli's publications, E. J. Stone proposed a new proof which was based upon two assumptions:

*Correction 4 (Stone 1873)*

A2

H7 The most probable value which can be adopted is that to which each individual measure equally contributes.

To obtain the most probable value, therefore, we must combine all the independent measures in such a way, that an error which may exist in one of the measures, as  $x_1$ , will produce the same error in the "value adopted as the most probable" as would be produced by the same error in  $x_2$ ,  $x_3$  or  $x_n$  suppose  $x_1, x_2, \dots x_n$  are  $n$  direct measures of the same quantity.

Let  $u = \phi(x_1, x_2, \dots x_n)$  be the value adopted as the most probable. Since equal errors (changes) in  $u$  are to be produced by the same error (change) in  $x_1, x_2, \dots x_n$  all partial differential derivations must be equal:

$$\frac{\partial u}{\partial x_1} = \frac{\partial u}{\partial x_2} = \dots = \frac{\partial u}{\partial x_n}$$

Schiaparelli had obtained the same equation in his first proof.

Stone shows that  $u$  is a function of the arithmetical mean  $s$ :  $u = F(s)$ . Then he tries to show by complete induction that  $F$  is the identity:

$F(s) = s$  (induction hypothesis)

Thus  $F'(s) = 1, F''(s) = F'''(s) = \dots = 0$

But Pizzetti (1892, 199) showed that this conclusion is not justified because  $u$  might depend on  $s$  as well as on  $n$ . Take for example

$$F(s) = s - \frac{n-2}{n} \left( s - \frac{a}{n} \log \left( 1 + \frac{ns}{a} \right) \right)$$

If  $n = 2$ , then  $F(s) = s$ , but this is not the case for any  $n \neq 2$ .

Joseph Marie de Tilly's article dating from 1874 was in a way the end of the series of attempts to prove the principle of arithmetical mean. He wanted to make evident the following four facts:

1. The principle of the mean for infinitely many values results immediately from the definition of accidental errors.
2. The principle can be proven for two quantities if we accept the postulate: There is an analytical function  $\phi$  of  $x$  which can represent the probability of an error between 0 and  $x$  (the probability of an error between  $x$  and  $x+dx$  can be expressed by  $\phi(x)dx$ ).
3. If we accept the principle for three quantities, it can be proven generally.
4. This is the last reduction which can be applied to it. Therefore, it is simpler to accept it completely from the very beginning.

De Tilly chose theorem 4 especially because the principle is practically useless in the two cases where it can be proved ( $n = 2, n = \infty$ ). None of his many predecessors had explicitly mentioned the postulate of theorem 2. But we need the

existence of such a function in order to know that the principle is right in the case  $n = 2$ . While Encke started with the "most reasonable hypothesis", i.e. that the errors are equal, de Tilly remarked that one should have assumed a priori the contrary.

In Pizzetti's opinion this assumption was one of the most severe flaws of error theory which was based upon the principle of the arithmetical mean: We suppose a priori that in general the relative frequency of an error depends on its magnitude; but it is an unjustified transition from this opinion to the assumption that there is a function which expresses this relative probability of an error for a given kind of observations.

After de Tilly, a certain opinion was more and more generally accepted. It was stated by Annibale Ferrero in (1876, 6) as follows:

"Whoever went back to another postulate applied one that was not more evident than the principle which was to be proved".

He admitted that analytical considerations always admit philosophical considerations in such problems: Though these trials of proving the principle did not achieve their purpose, they revealed the open or hidden presuppositions which underlay the application of the principle of error theory. Indeed, in spite of certain similarities between these described proofs and refutations, the principle of the arithmetical mean is not another example of Imre Lakatos' methodology of mathematical development. All these efforts came to nothing. The principle was an empirical rule that asserted that the arithmetic mean is the "best" estimator of the standard linear model without defining the optimality criterion. The authors tried to show, by appropriate systems of axioms, that the arithmetical mean is the only permissible estimator of this model (Farebrother 1985).

#### *4.2. The Second Alternative or the Role of Experience*

Up to now we discussed the first alternative mentioned by Encke with regard to the principle. But the second alternative was not less problematical. For the principle of the arithmetical mean is not universally valid. Already in 1760 J. H. Lambert mentioned in his *Photometria* two examples where the arithmetical mean did not seem to give the greatest approximation to the truth (Lambert 1760, § 276, 277).

##### *First Example:*

If we consider the perimeter of an  $n$ -gon as an observation of the length of the perimeter of the circle, then the arithmetical mean of the inscribed and the circumscribed  $n$ -gon does not supply the most probable value of its perimeter.

Both Encke (1834, 263) and Glaisher (1872, 91 f) commented upon this

example, but in different ways. Encke accepted this example as such. But in his opinion these observations are not equally good. Thus Lambert's example cannot impair the validity of the principle which presupposes this condition. Glaisher rejected this example. The lengths of the perimeters are by their nature no observations and consequently no quantities the principle of the arithmetical mean might be applied to.

Thus even the concept of observation which leads to errors had become problematical because the nature of errors had not sufficiently been clarified up to then. In Glaisher's opinion each actual error is found by the linear combination of a large number of errors due to different independent sources. "This supposition not only seems to be a true one, but also to include all that can be asserted with anything approaching to certainty of the nature of an error". In other words, the philosophical basis (hypothesis of elementary errors) was crucial for accepting or rejecting counterexamples against the mathematical notion, proof, or method.

*Second Example:*

Neither Encke nor Glaisher mentioned Lambert's observation that we have to take the geometrical mean in the case of ratios, as for the example in the case of photometric measurements. In 1834 Ernst Heinrich Weber had inquired into the psychophysical law which is called the law of Weber and Gustav Theodor Fechner. Two years later Bessel's pupil von E. A. Steinheil showed that the difference between the sizes of stars are approximately proportional to the differences between the logarithms of their light intensity (Steinheil 1836).

Therefore, von Seidel called the logarithm of the ratio of the light intensities of two stars, the difference between the light intensities. For the different logarithms are subject to equally probable errors. The analogous assumption that the same is true with the numbers themselves would be absurd (Seidel 1863, 426).

For unknown reasons it was several decades before Seidel's article was taken notice of. The reactions were again very different. Pizzetti denied that the law of Weber and Fechner might be applied to the usual astronomical researches and to error theory because an error of observation in its entirety depends only to a minimal part on the defects of human senses (1892, 207 f). H. Seeliger (1893), on the contrary, underlined that the arithmetical mean is far from always supplying the most probable value, explicitly referring to Seidel. Seeliger conceived a mathematical reason: It is an inadmissible assumption that the relative frequency of an error  $x-1$  depends on itself, but no longer on the most probable value  $x$  of the unknown quantity or on the observed value 1.

Thus there is a double difficulty in applying the principle of the arithmetical mean: (i) We have to make sure that it supplies the most probable value.

(ii) We have to prove this statement. It does not suffice to prove that it supplies only the most plausible value as Reuschle and many other authors intended to do because the method of least squares proceeds from the most probable value. The most plausible value is worthless for the foundation of the method of least squares. In other words: the philosophical basis (mathematical description of the nature of errors) called for a mathematical proof that certain presuppositions were fulfilled, thus deciding whether a principle was admissible.

Were all these controversial attempts of founding error theory in vain? The answer is certainly no. We can say that the constructive aspect of such diverging approaches to error theory laid the foundations of the modern theory of invariant tests and estimators (Sheynin 1979, 32).

### References

- Adrain, R. (1808), "Research concerning the probabilities of the errors which happen in making observations, etc.", *The Analyst or Mathematical Museum* 1(4), 93–109. Reprinted in (Stigler 1980, I).
- Airy, G. B. (1861), *On the algebraical and numerical theory of errors of observations and the combination of observations*, London, Macmillan.
- Bertrand, J. L. F. (1889), *Calcul des probabilités*, Paris, Gauthier-Villars. All page references are to the second edition Paris, 1907.
- Bessel, F. W. (1838), "Untersuchungen über die Wahrscheinlichkeit der Beobachtungsfehler", *Astronomische Nachrichten* 15, 369–404.
- Bienaymé, I. J. (1852), "Sur la probabilité des erreurs d'après la méthode des moindres carrés", *Journal de Mathématiques pures et appliquées* (1) 17, 33–78.
- Bienaymé, I. J. (1853), "Considérations à l'appui de la découverte de Laplace sur la loi de probabilité dans la méthode des moindres carrés", *Comptes rendus hebdomadaires des Séances de l'Académie des Sciences* 37, 309–324.
- Chabert, J.-L. (1989), "Gauss et la méthode des moindres carrés", *Revue d'Histoire des Sciences* 42, 5–26.
- Chauvenet, W. (1868), *A manual of spherical and practical astronomy: embracing the general problems of spherical astronomy. The special applications to nautical astronomy, and the theory and use of fixed and portable astronomical instruments, with an appendix on the method of least squares*. Vol. II, *Theory and use of astronomical instruments, Method of least squares*, Philadelphia, J. B. Lippincott.
- Crofton, M. W. (1870), "On the proof of the law of errors of observations", *Philosophical Transactions of the Royal Society of London* 160, 175–187.
- Czuber, E. (1899), "Die Entwicklung der Wahrscheinlichkeitsrechnung und ihrer Anwendungen", *Jahresbericht der Deutschen Mathematiker-Vereinigung* 7 (1898), Teil 2, Leipzig.
- Dienger, J. (1852), "Über die Ausgleichung der Beobachtungsfehler", *Archiv für Mathematik und Physik* 18, 149–193; 19 (1852), 211–227.
- Donkin, W. F. (1851), "On certain questions relating to the theory of proba-



- bilities, Part III", *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science* (4) 1, 55–60.
- Donkin, W. F. (1857), "On an analogy relating to the theory of probabilities of least squares", *Quarterly Journal of pure and applied mathematics* 1, 152–162.
- Edgeworth, F. Y. (1883), "The method of least squares", *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* (5) 16, 360–375.
- Edgeworth, F. Y. (1884), "A priori probabilities", *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* (5) 18, 204–210.
- Ellis, R. L. (1844), "On the method of least squares", *Transactions of the Cambridge Philosophical Society* 8, 204–219.
- Ellis, R. L. (1850), "Remarks on an alleged proof of the 'Method of least squares', contained in a late Number of the Edinburgh Review. In a Letter addressed to Professor J. D. Forbes", *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* (3) 37, 321–328 with a supplement on p. 462.
- Encke, J. F. (1832), "Über die Begründung der Methode der kleinsten Quadrate", *Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin, Aus dem Jahre 1831, Mathematische Klasse*, Berlin, 73–78.
- Encke, J. F. (1844), "Bemerkungen zu der Abhandlung No. 22. Band 26. Heft 4. dieses Journals, (=Reuschle 1843)", *Journal für die reine und angewandte Mathematik* 28, 213–222.
- Farebrother, R. W. (1985), "The statistical estimation of the standard linear model, 1755–1853", *Proceedings of the first International Tampere Seminar on Linear Statistical Models and Their Applications*, Tampere, August 1983, Tampere, University of Tampere, 77ff.
- Ferrero, A. (1876), *Esposizione del metodo dei minimi quadrati*, Florence, G. Barberá.
- Gauss, C. F. (1809a), *Theoria motus corporum coelestium*, Hamburg, F. Perthes u. I. H. Besser. Reprinted in (Gauss 1906, 1–288). All page references are to the German selected translation (Gauss 1887, 92–117).
- Gauss, C. F. (1809b), "Self-announcement of (Gauss 1809a)", *Göttingische Gelehrte Anzeigen*. Reprinted in (Gauss 1874, 53–60). All page references are to (Gauss 1887, 204 f).
- Gauss, C. F. (1821/23), "Self-announcement of (Gauss 1823)", *Göttingische Gelehrte Anzeigen* Stück 33 (1821), 321–327; Stück 32 (1823), 313–318. Reprinted in (Gauss 1873, 95–104). All page references are to (Gauss 1887, 190–199).
- Gauss, C. F. (1823), "Theoria combinationis observationum erroribus minimis obnoxiae" (2 parts: 1821, 1823), *Commentationes societatis Regiae scientiarum Göttingensis recentiores* 5, 33–90. Reprinted in (Gauss 1873, 1–53). All page references are to the German selected translation (Gauss 1887, 1–53).
- Gauss, C. F. (1862), *Briefwechsel zwischen C. F. Gauss und H. C. Schumacher*, ed. by C. A. F. Peters, vol. 4, Altona, Gustav Esch. Reprinted in C. F. Gauss, *Werke, Ergänzungsreihe*, Bd. 5, *Briefwechsel C. F. Gauss - H. C. Schumacher*, Teil 2, Hildesheim-New York, Olms, 1975.
- Gauss, C. F. (1873), *Werke*, vol. 4, Göttingen, Dieterichsche Universitäts-Buch-

- druckerei. All page references are to the second edition, Göttingen, 1880.
- Gauss, C. F. (1874), *Werke*, vol. 6, Göttingen, Dieterich.
- Gauss, C. F. (1906), *Werke*, vol. 7, Leipzig, Teubner.
- Gauss, C. F. (1880), *Werke, Ergänzungsreihe* Bd. 1, *Briefwechsel C. F. Gauss - F. W. Bessel*, Leipzig, Wilhelm Engelmann. Reprint, Hildesheim-New York, Olms, 1975.
- Gauss, C. F. (1887), *Abhandlungen zur Methode der kleinsten Quadrate von Carl Friedrich Gauss*, ed. by A. Börsch and P. Simon, Berlin. All page references are to the reprint Würzburg, Physica, 1964.
- Glaisher, J. W. L. (1872), "On the law of facility of errors of observations, and on the method of least squares", *Memoirs of the Royal Astronomical Society* 39, 75-124.
- Hagen, G. H. L. (1837), *Grundzüge der Wahrscheinlichkeitsrechnung*, Berlin, Dümmler. All page references are to the second edition, Berlin, Ernst & Korn, 1867.
- Herschel, J. (1850), "Review of A. Quetelet, *Lettres à S. A. R. le Duc régnant de Saxe-Cobourg et Gotha sur la théorie des probabilités appliquée aux Sciences Morales et Politiques*, Brüssel 1846", *The Edinburgh Review, or Critical Journal* 92, 1-30.
- Ivory, J. (1825), "On the method of least squares", *The Philosophical Magazine and Journal comprehending the various branches of science, the Liberal and Fine Arts, Agriculture, Manufactures and Commerce* 65, 3-10, 81-88, 161-168.
- Ivory, J. (1826), "On the method of least squares", *The Philosophical Magazine and Journal comprehending the various branches of science, the Liberal and Fine Arts, Agriculture, Manufactures and Commerce* 68, 161-165.
- Knobloch, E. (1985), "Zur Grundlagenproblematik der Fehlertheorie". In: *Mathemata, Festschrift für Helmuth Gericke*, ed. by M. Folkerts and U. Lindgren, Stuttgart, Steiner, 561-590.
- Lambert, J. H. (1760), *Photometria*, Augsburg, Klett.
- de Laplace, P. S. M. (1810), "Mémoire sur les approximations des formules qui sont fonctions de très grands nombres et sur leur application aux probabilités", *Mémoires de l'Académie des Sciences de Paris*, 1809, 353-415, 559-565. Reprinted in *Oeuvres*, vol. 12, 301-353.
- de Laplace, P. S. M. (1812), *Théorie analytique des probabilités*, Paris, Courcier. All page references are to the 3rd edition, Paris 1820 (with supplements). Reprinted as *Oeuvres*, vol. 7, Paris, 1886.
- de Laplace, P. S. M. (1816), "Premier supplément" (to the *Théorie analytique des probabilités*). All page references are to *Oeuvres*, vol. 7, Paris, 1886.
- Legendre, A.-M. (1805), *Nouvelles méthodes pour la détermination des orbites des comètes*, Paris, Firmin Didot.
- Lindelöf, L. (1871), "Review of V. Neovius, *Lärobok i minsta quadrat-metoden*, Åbo, 1870", *Bulletin des sciences mathématiques et astronomiques* 2, 134-136.
- Merriman, M. (1877a), "A list of writings relating to the method of least squares, with historical and critical notes", *Transactions of the Connecticut Academy of Arts and Sciences* 4, 151-227. All page references are to (Stigler 1980, I).

- Merriman, M. (1877b), "On the history of the method of least squares", *The Analyst* 4 (4 pages). All page references are to (Stigler 1980, I).
- de Morgan, A. (1864), "On the theory of errors of observation", *Transactions of the Cambridge Philosophical Society* 10, 409–427.
- Peirce, C. S. (1870), "On the theory of errors of observations", *Report of the U. S. Coast and Geodetic Survey*, App. 21.
- Pizzetti, P. (1892), "I fondamenti matematici per la critica dei risultati sperimentali", *Atti della Regia Università di Genova*, 113–333.
- Reuschle, K. G. (1843), "Über die Deduction der Methode der kleinsten Quadrate aus Begriffen der Wahrscheinlichkeitsrechnung", *Journal für die reine und angewandte Mathematik* 26, 333–364.
- Reuschle, K. G. (1844), "Zusätze zu der Abhandlung über die Methode der kleinsten Quadrate", *Journal für die reine und angewandte Mathematik* 27, 182–184.
- Schiaparelli, G. V. (1868), "Sul principio della media aritmetica nel calcolo dei risultati delle osservazioni", *Reale istituto Lombardo di scienze e lettere, Rendiconti* (2) 1, Milano, 771–778.
- Schneider, I. (1981), "Die Arbeiten von Carl Friedrich Gauss im Rahmen der Wahrscheinlichkeitsrechnung: Methode der kleinsten Quadrate und Versicherungswesen". In: *Carl Friedrich Gauss (1777–1855), Sammelband von Beiträgen zum 200. Geburtstag von C. F. Gauß*, ed. by I. Schneider, München, Minerva, 143–172.
- Schneider, I. (Ed.) (1988), *Die Entwicklung der Wahrscheinlichkeitstheorie von den Anfängen bis 1933. Einführungen und Texte*, Darmstadt, Wissenschaftliche Buchgesellschaft.
- v. Seeliger, H. (1893), "Bemerkung über das arithmetische Mittel", *Astronomische Nachrichten* 132, col. 208–214.
- v. Seidel, Ph. L. (1863), "Resultate photometrischer Messungen an 208 der vorzüglichsten Fixsterne", *Abhandlungen der Königlich Bayerischen Akademie der Wissenschaften, Mathematisch-physikalische Klasse* 9, 3. Abt., 419–609.
- Sheynin, O. B. (1979), "C. F. Gauss and the theory of errors", *Archive for History of Exact Sciences* 20, 21–72.
- v. Steinheil, C. A. (1836), *Elemente der Helligkeitsmessungen*, München (Denkschriften der Königlich Bayerischen Akademie der Wissenschaften, Klasse II).
- Stigler, St. M. (Ed.) (1980), *American contributions to mathematical statistics in the nineteenth century*, 2 vols., New York, Arno Press. (These volumes do not contain an own pagination).
- Stigler, St. M. (1981), "Gauss and the invention of least squares", *Annals of Statistics* 9, 465–474.
- Stigler, St. M. (1986), *The history of statistics. The measurement of uncertainty before 1900*, Cambridge (Mass.)-London, Harvard UP.
- Stone, E. J. (1873), "On the most probable result which can be derived from a number of direct determinations of assumed equal value", *Monthly notices of the Royal Astronomical Society* 33, 570–572.
- Tait, P. G. (1865), "On the law of frequency of error", *Transactions of the Royal Society of Edinburgh* 24, 139–145.

- de Tilly, J. M. (1874), "Note sur le principe de la moyenne arithmétique, et sur son application à la théorie mathématique des erreurs", *Nouvelle correspondance mathématique* 1, 137–147.
- v. d. Waerden, B. L. (1977), "Über die Methode der kleinsten Quadrate", *Nachrichten der Akademie der Wissenschaften in Göttingen, Mathematisch-Physikalische Klasse*, 75–87.
- Young, Th. (1819), "Remarks on the probabilities of error in physical observations, and on the density of the earth, considered, especially with regard to the reduction of experiments on the pendulum. In a letter to Capt. Henry Kater, F. R. S.", *Philosophical Transactions of the Royal Society of London*, 70–95.

# A Structuralist View of Lagrange's Algebraic Analysis and the German Combinatorial School

HANS NIELS JAHNKE (Bielefeld)

## *1. Motivations \**

Two observations motivate the following considerations pertaining to the structure of mathematical theories at the beginning of the 19th century. One is related to the history of mathematics, and the other to its philosophy.

It is generally believed that Kantian philosophy of mathematics exerted a strong influence in Germany during the 19th century and that it underwent a crisis only in connection with the development and elaboration of non-Euclidean geometries. This view is contradicted by some observations which show that, from its beginning, Kant's philosophy did not encompass the key trends of contemporary mathematics. Kant had defined mathematics as knowledge whose subject was our a priori (pure) intuition of space and time. More specifically, he had said that mathematics is rational cognition based on the construction of concepts in space and time.<sup>1</sup> True, his notion of a "symbolical construction" extended to algebra, but his philosophy could be readily applied only to geometry and elementary arithmetic. The analytical calculus, the most important part of contemporary mathematics, played no explicit role in his philosophy. This had been criticized already in his time. Thus, the famous linguist J. G. Herder remarked that Kant's philosophy of mathematics rested on the "radical misconception...[as if] visible construction could exhaust the matter".<sup>2</sup> O. Becker, in the twenties of our century, spoke of a "classicist, one might say, reactionary turn" in Kant.<sup>3</sup>

How remote from Kant's intellectual world mathematics had become at the

---

\* I appreciate gratefully the kind support of Abe Shenitzer and Hardy Grant, Toronto, in producing a readable English version of this paper.

1 Kant [1781], B 741/2.

2 Herder [1799], 265.

3 Becker [1927], 297.

turn of the 19th century is shown by the fact that textbooks of the time used a classification of mathematics in which only arithmetic (including algebra and infinitesimal analysis) was included in pure mathematics while geometry and the branches of theoretical physics (mechanics and astronomy) were seen as parts of applied mathematics. For example, consider the following classification due to the German mathematician M. Ohm (1792–1872):<sup>4</sup>

- I. Arithmetic (= pure mathematics)
  - 1. Elementary arithmetic
  - 2. Higher arithmetic (= higher algebra and analysis)
- II. Theory of quantities (= applied mathematics)
  - 1. General theory of quantities
  - 2. Special theory of quantities
    - i) Theory of geometric quantities (= geometry)
    - ii) Theory of mechanical quantities (= mechanics)

This picture of the structure of mathematics is also mirrored in a famous remark of Gauss. "It is my innermost conviction that the relation of geometry to our knowledge a priori is radically different from that of the pure theory of magnitudes; our knowledge of the former completely lacks that sense of necessity (and thus of absolute truth) which is inherent in the latter; we must confess with humility that whereas number is solely the product of our mind, space has a reality beyond our mind – a reality whose laws we cannot completely prescribe a priori".<sup>5</sup> Although this remark was motivated by Gauss' ideas about non-Euclidean geometry, it also reflected widespread views which had developed independently of non-Euclidean geometry.

This change of views from Kant's time to the early 19th century has not yet been thoroughly analyzed. I propose to show that this change is closely related to the rise and fall of the so-called Combinatorial School in Germany whose members had attempted to build all of analysis on an algebraic-combinatorial basis. This approach was linked to attempts by other mathematicians, above all J. L. Lagrange in his *Théorie des fonctions analytiques* (1797) to treat infinitesimal analysis in a purely algebraic manner. Following the triumph of Cauchy's conception of infinitesimal analysis, these algebraic approaches were considered complete failures, because they seemed to imply an inherent circularity, presumably overlooked by their authors. The criticism of Lagrange's attempt by the German mathematician Hermann Hankel was withering: "... and this natural method he (Lagrange) found in the definition of

<sup>4</sup> Ohm [1822]. Cf. Bekemeier [1987].

<sup>5</sup> Letter of Gauß to Bessel, 9, April 1830 (Becker[1975], 179).

the differential quotient  $f'(x)$  of  $f(x)$  as the coefficient of  $h$  in the expansion of  $f(x+h)$  in powers of  $h$ , whereby he believed to have eliminated from analysis the notions of the infinite and of limit. But the famous mathematician overlooked the fact that his definition stipulated in an essential way the infinite nature of the series for  $f(x+h)$ , and that the notion of infinite series necessarily implied the infinite and a notion of limit, and therefore suffered from exactly the same shortcomings as the notions of the differential quotient and of the integral. To us, who, from our youth are used to the rigorous notion of a convergent series, it is inconceivable that the brilliant author of the *Théorie des fonctions*, whose favorite idea was to restore to science the rigor of the ancients, could overlook this obvious requirement. For those who would learn from history, this could be a warning example that even the mathematician cannot withdraw with impunity from the ground of the natural and of the historically actualized". Hankel judged Lagrange's attempt to found analysis to have been the "poorest of all".<sup>6</sup>

If correct, the charge of a vicious circle would have been equally justified in the case of the combinatorial school. We will examine this charge by looking at some of the textbooks of this group and will try to show that, from a philosophical viewpoint, the charge is false. It will become clear that Lagrange and the combinatorial school had in mind a rather modern view of a theory that consistently stressed the difference between its pure and applied components. From a mathematical point of view, Cauchy was not superior to Lagrange because his theory was free of circularity but because his approach made possible finer distinctions. In a logical sense, Lagrange and the combinatorial school provided a legitimate alternative to Cauchy, whereas, from a philosophical viewpoint, it could be argued that they provided a preferable alternative to Kant's philosophy of mathematics.

## 2. Theory and Applications

To get an idea of the structure of mathematical theories envisaged by the combinatorial school, we will analyze a textbook from the twenties of the 19th century that is typical for this period and that treats in detail the combinatorial approach to infinitesimal analysis.<sup>7</sup>

This textbook belongs to the "second period" of the combinatorial school, when this group had lost its dominant position in German mathematics but when a number of textbooks appeared in which the combinatorial approach

---

<sup>6</sup> Hankel [1871], 200, 209.

<sup>7</sup> Spehr [1826].

was presented very lucidly.<sup>8</sup> Its author, Friedrich Wilhelm Spehr, was born in 1799 as the son of a merchant. Beginning in 1819, he studied mathematics in Göttingen under Gauss and B. F. Thibaut. In 1824 he published a treatise on combinatorics which very quickly made him famous and helped him to secure a position at the Collegium Carolinum in Braunschweig (where R. Dedekind was to teach later). The exertions of surveying the dukedom of Braunschweig and private problems, undermined his health. He died in 1833 at the age of 34.

From the longwinded title of Spehr's textbook we learn that its author intended to present the differential and the variational calculus *independently of the usual method of fluxions, of the concepts of the infinitely small and of vanishing magnitudes, of the method of limits, and of the theory of functions*. According to Spehr's ideas, this objective was to be achieved by consistently distinguishing between the pure and the applied parts of analysis. The general distinction between pure (a priori) and applied (a posteriori) science derived from Kant and played a considerable role in the scientific literature of the time. This distinction was presumably the key to the conceptual control of any subject and to the elimination of hidden assumptions. This minimized the conceptual apparatus of a theory and allowed one to gain greater insight into its empirical content. Thus, in drawing a distinction between pure and applied analysis, Spehr followed a then current usage.

Applied to analysis, this distinction acquires a specific and somewhat surprising meaning. Spehr distinguished between "analysis" as a pure science and "differential and integral calculus" as an applied science. According to Spehr, "Analysis is the science of the laws of combination of composite numbers, and its main object is the function". For him, "function" meant a symbolical expression. "A function of one or several principal magnitudes [*Hauptgrößen*] is thus an expression arithmetically composed in some manner from those principal magnitudes and other, secondary, magnitudes [*Nebengrößen*]"<sup>9</sup> He then added the footnote:

"Although we intend gradually to substitute definite magnitudes for the principal magnitudes of a function, whereby the function takes on ever different values, it is not appropriate to call the principal magnitudes and the function itself variable magnitudes, because this concept must be reserved for the differential calculus where we first encounter true variability [in the form of] the flowing magnitude. In that science, in which one imagines – or should imagine – that the flowing magnitude passes according to a definite law through all states it can attain in accordance with this law, one should bear in mind that the originally variable magnitudes that conform to the

---

<sup>8</sup> cf. Jahnke [1990], chap. III.

<sup>9</sup> Spehr [1824], 141.



principal magnitudes in analysis take on every conceivable positive and negative value, whereas the substitutions involving the principal magnitudes in analysis are, for the most part, very limited. But the chief aim of all analytical investigations is the subsequent application to the differential calculus; for the most part, the operations with functions are dealt with for the sole reason of using them afterwards in the differential calculus, that is, in order to view the principal magnitudes, including the functions, as variable magnitudes".<sup>10</sup>

This means that a clear distinction is drawn between the formal theory, called analysis, which investigates the arithmetical (= algebraic) composition of principal and secondary magnitudes, and the material theory, called differential and integral calculus, which essentially has as its subject the concepts of variability and of continuous magnitude. This was a basic distinction of the combinatorial school from its inception. Thus, algebraic (or combinatorial) analysis, that is analysis of the finite, as treated, say, in L. Euler's *Introductio in analysin infinitorum*,<sup>11</sup> is analysis in the proper sense of the term, whereas differential and integral calculus, that is analysis of the infinite, is just an application of algebraic analysis. C. F. Hindenburg (1741–1808), the *spiritus rector* of the combinatorial school, had written: "... in general, analysis proper investigates the forms of magnitudes. From this two things follow very readily: one is the great usefulness of the combinatorial method, whose main objective is the expansion, representation, and study of such forms, and the other is its immediate applicability".<sup>12</sup> Thus, analysis proper is concerned only with the study of forms, that is, of transformations of finite and infinite symbolical expressions, above all of formal power series.

The so-called polynomial theorem played a key role in the calculus of formal power series. This theorem states that the  $m$ -th power of an arbitrary polynomial or infinitinomial (*infinitorum*), that is, of a power series, is again a power series,

$$(1 + a_1 x^1 + a_2 x^2 + \dots)^m = 1 + A_1 x^1 + A_2 x^2 + A_3 x^3 + \dots,$$

whose coefficients can be calculated from the formula

$$A_r = \sum_{h=1}^r \binom{m}{h} p^h {}^h C_r;$$

here  $p$  as operator denotes the appropriate polynomial coefficient and the  ${}^h C_r$  are combinatorially built up from the coefficients  $a_i$  in an obvious manner.

It is clear that the right-hand side of the equation can be formally

<sup>10</sup> l.c., 138.

<sup>11</sup> Euler [1748].

<sup>12</sup> Hindenburg [1796], "Vorbericht".

interpreted even in the case of rational or negative exponents  $m$ . This is so because  $m$  appears only in the binomial coefficients and because every  $A_r$  is calculated from a finite sum.

This general polynomial theorem is of central importance for all calculations with power series, because the operations of division (using negative exponents), of exponentiation and of extracting an arbitrary root (using fractional exponents) can be reduced to this formula. Moreover, in 1793, H. A. Rothe (1773–1842), one of Hindenburg's disciples, showed that an arbitrary algebraic or transcendental equation expressed in terms of formal series,

$$a_1x^1 + a_2x^2 + a_3x^3 + \dots = b_1y^1 + b_2y^2 + b_3y^3 + \dots,$$

can be solved for  $x$  or  $y$  by applying the polynomial formula.<sup>13</sup> This theorem about the reversion of series was a kind of implicit function theorem for power series. It was deemed a remarkable success, all the more so because Rothe and J. F. Pfaff (1765–1825) were able to show two years later that this formula is equivalent to a famous theorem of Lagrange which effects the reversion of series by methods of the differential calculus.<sup>14</sup> Thus, for Hindenburg and his adherents, the polynomial formula was the most important theorem of analysis.<sup>15</sup>

Euler had proved the polynomial theorem by means of the calculus.<sup>16</sup> To show that analysis as a universe of symbolic expressions in the sense of the combinatorial school can be built up in a purely algebraic-combinatorial way, it was crucial to show that the polynomial theorem can be proved by purely algebraic-combinatorial means.<sup>17</sup> A combinatorial proof called for a purely formal interpretation of the polynomial formula that was independent of the convergence or divergence of the series involved. While these ideas were never fully elaborated, this approach imparted an algebraic identity not only to analysis but also to its applications, that is, to the differential and integral calculus.

The structure of analysis is consistently elaborated in Spehr's textbook. For him analysis is a theory of symbolic expressions, while the differential

---

<sup>13</sup> Rothe [1793]. An algorithmic procedure for reverting power series had already been given by Newton (Newton [1676]), and the technique played an important part in Newton's work on calculus.

<sup>14</sup> cf. Lagrange [1770], Pfaff [1795a], Pfaff [1795b], Rothe [1795].

<sup>15</sup> cf. Hindenburg [1796].

<sup>16</sup> Euler [1755], pars posterior, § 202.

<sup>17</sup> cf. the title of the paper Gudermann [1825]. Gudermann later became the academic teacher of Weierstraß. The paper was again published in 1830 in *Crelle's Journal*.

and integral calculus are mere applications of this pure theory. In an introduction of 120 pages he treated those subjects which one would expect to find in a modern textbook of infinitesimal analysis, such as Taylor's theorem, rules of differentiation and integration, and differential equations. The treatment always focusses on the arithmetical and formal aspects of the theory. This is followed by an account of the calculus of fluents (*Fluentenkalkül*), comprising the laws of continuous magnitudes. The second part of the book contains a general account of continuous magnitudes followed by applications to geometry and mechanics. Spehr justified his approach with the remark that earlier authors had not sufficiently separated the purely arithmetical investigations from the differential calculus and that some authors, such as Lagrange, had thought that the arithmetical investigations were the essence of the differential calculus. But the essentially new concept of a continuous magnitude made it impossible to reduce the differential calculus to arithmetical calculations. He had therefore rigorously separated the formal parts of the theory from its material ones and had established the concept of a continuous magnitude as the basic concept of the material part of the subject.<sup>18</sup>

### 3. Taylor's Theorem

To gain a deeper insight into the structure of Spehr's theory we will discuss two issues, namely, Taylor's theorem as a central topic within the formal theory (*analysis*) and the definition of continuity as a basic concept of the material theory (*differential calculus*).

We begin with the formal part. In Spehr's view Taylor's theorem is a purely syntactical transformation. He started his proof of Taylor's theorem by noting that if  $\varphi$  is a function of  $x$  and if  $x$  in  $\varphi(x)$  is replaced by  $x + h$ , then, as shown in analysis,  $\varphi(x+h)$  can be expanded in a series of successive powers of  $h$ ;

$$\varphi(x+h) = \overset{0}{A} + \overset{1}{A} h^1 + \overset{2}{A} h^2 + \dots + \overset{r}{A} h^r + \dots$$

From this relation and from an equation for the  $k$ -th difference of the functional values  $y$  he calculates the coefficients  $\overset{k}{A}$ ,

$$\overset{k}{A} = \frac{\overset{1}{T} \Delta^k y}{1 \cdot 2 \cdot \dots \cdot k \cdot \Delta x^k}$$

Here  $\overset{1}{T} \Delta^k y$  denotes the first term of the series expansion of  $\Delta^k y$ , the  $k$ -th differences of  $y$ , in powers of  $\Delta x$ .

---

<sup>18</sup> Spehr [1826], VIII–XI.

In this way, Spehr obtains the following expansion:

$$\Delta y = \overset{1}{T} \Delta y + \frac{\overset{1}{T} \Delta^2 y}{1 \cdot 2} + \frac{\overset{1}{T} \Delta^3 y}{1 \cdot 2 \cdot 3} + \dots + \frac{\overset{1}{T} \Delta^k y}{1 \cdot 2 \cdot \dots \cdot k} + \dots$$

Then he defines  $d^k y$ , the  $k$ -th order differential of  $y$ , by means of the equation

$$\overset{1}{T} \Delta^k y = d^k y,$$

and thus obtains the Taylor series

$$\Delta y = dy + \frac{d^2 y}{1 \cdot 2} + \frac{d^3 y}{1 \cdot 2 \cdot 3} + \frac{d^k y}{1 \cdot 2 \cdot \dots \cdot k} \dots$$

"This series progresses according to powers of  $\Delta x$ , and  $dy$  contains  $\Delta x$ ,  $d^2 y$  contains  $\Delta x^2$ , and, in general,  $d^k y$  contains  $\Delta x^k$ ".<sup>19</sup>

From the viewpoint of formal series, the proof is correct. This is so because only finitely many power series in  $\Delta x$  are summed, so that the respective coefficients of  $(\Delta x)^k$  in the last equation are calculated from a finite sum.

In this proof, no infinite quantities or limits are mentioned, and, according to modern standards, the whole procedure is a mere tautology, because it assumes what is to be proved, namely the possibility of expanding a function in a power series. But this criticism is irrelevant. Spehr and other contemporary authors proved a different theorem from that proved today. Spehr's theorem states that a certain transformation can be performed by means of a general algorithm. If a function  $f$  can be expanded in a series, then the coefficients of this series representation can be calculated by means of a certain algorithm. To obtain the first coefficient of the series we must transform the difference  $f(x + \Delta x) - f(x)$  into a product form  $f'(x) \cdot \Delta x$ . The next coefficient is obtained by applying the same operation to  $f'(x)$ , and so on. In some cases, this transformation is possible, and in others it is not. The question of whether the formal relations calculated in this way lead to correct numerical equations depends on the particular function  $f$ . In other words, we can interpret Spehr's introductory claim that it is proved in analysis that every function can be expanded in a series to mean that in every individual case we must investigate whether or not the function involved belongs to the functions of analysis.

The modern version of Taylor's theorem asserts that under certain conditions (including uniform convergence), two different procedures for the calculation of numerical values, i.e., the functional law and Taylor's series yield the same numerical values. This shows that both versions of the theorem have the same logical structure. In both cases, in Spehr's version and in the modern

---

<sup>19</sup> Spehr [1826], 13–17.

one, one must determine whether or not the conditions for the application of Taylor's theorem are fulfilled.

This interpretation of Spehr's version of Taylor's theorem is not an attempt to "save" a basically circular and tautological proof. Rather, it is an accurate reflection of the contemporary practice of applying Taylor's theorem. In particular, this interpretation catches the essence of Lagrange's procedure in his *Théorie des fonctions analytiques*. Lagrange, too, begins with the assumption that a function can be expanded in a series.<sup>20</sup> If this is not to be an absurd statement – and we shouldn't impute an absurdity to such an eminent mathematician – it can only mean that this assumption has already been verified for the individual functions and classes of functions which are to be considered.<sup>21</sup> This interpretation is all the more plausible because Lagrange states on various occasions that there are functions which cannot be expanded in a series and that there are singular points where such an expansion does not hold.<sup>22</sup>

Lagrange continues by deriving Taylor's formula. For him this formula provides an algorithm for the concrete calculation of series expansions. Contrary to Hankel's assertion, Lagrange's procedure is far from circular. Lagrange had not overlooked – as Hankel maintains – the need for convergence. Rather, he assigned to this concept a particular role. It was Hankel who seems to have underestimated the fact that we owe to Lagrange the Lagrange remainder formula,<sup>23</sup> one of the most important techniques for calculating the (numerical) error when approximating a function numerically by its Taylor expansion. Also, it is not true, that the concept of convergence, although intentionally excluded, is introduced into the theory through the backdoor, as Hankel would have us believe. Rather, Lagrange's approach may be described as follows.

Lagrange's idea was to define the concepts of derivative and integral independently of limit ideas and of notions of infinitely small magnitudes, and that meant for him, independently of geometric intuition and of numerical approximations. He achieved this in a formally elegant way by defining the derivative of a function as the coefficient of  $h$  in the series expansion of  $f(x+h)$ . If this expansion is considered as a purely formal operation which does not presuppose convergence or divergence, then it is, in fact, independent of convergence, and Lagrange's scheme is free of circularity. The question of convergence arises only in the second step, the step of numerical evaluation,

---

<sup>20</sup> Lagrange [1797], I, §1, §2, §7.

<sup>21</sup> In §2 it is explicitly stated that the assumption is justified by the expansion of the known functions.

<sup>22</sup> l.c.

<sup>23</sup> l.c., § 33–40.

which is conceptually independent of the calculus of formal power series. It bears repeating that far from being unaware of this problem, Lagrange has provided important new techniques of numerical evaluation, such as his remainder theorem, and techniques for solving equations described in his fundamental *De la résolution des équations numériques de tous les degrés*.<sup>24</sup>

We do not claim that Lagrange had worked out all details of his approach in a mathematically satisfactory way. We know that this was not the case. But it is our intention to show that Lagrange (and the combinatorial school) had at least a vision of a theory that was free of circularity (but had some gaps), and that could be fully elaborated.<sup>25</sup> The inner logic of this conception had important advantages. First, the concept of a derivative was given by a fully abstract and theoretical definition, which was independent of any "empirical" idea of numerical approximation. Second, analysis was considered as something that is conceptually more general than its particular applications to geometry and mechanics. Third, it separated the theory from its applications.

#### 4. *The Dichotomy of the Continuous and the Discontinuous within the Combinatorial Approach*

We come back to Spehr's textbook and ask how he conceived of the applied, material part of analysis, i.e., of the differential and integral calculus. What did it mean to make the concept of a continuous magnitude the fundamental concept of the differential and integral calculus? As we said, the arithmetical part of the theory was in place before this concept was introduced, so that the continuity concept could contribute nothing to it. With Spehr, the concept of continuity did not belong to analysis, but to the applied disciplines of infinitesimal analysis – geometry and mechanics. Hence, he saw no need to formalize this concept mathematically. Instead, he gave a purely verbal definition of continuity and of other related concepts. "A continuous magnitude, a continuum, is any magnitude which is thought to be in a state of becoming, that advances not by leaps, but by uninterrupted progress. Thus, any arbitrary curve results from the movement of a point in space, a surface from the

---

<sup>24</sup> Lagrange [1798].

<sup>25</sup> To be more precise: J. V. Grabiner in her [1981] uncovered the roots of Cauchy's techniques of proof in 18th century analysis, especially in the work of Lagrange. This showed, that Lagrange's position was much nearer to Cauchy's than is usually thought. Here, I agree with Grabiner. On the other hand, I am inclined to impute to Lagrange, in the light of later works like those of Spehr and M. Ohm, a much clearer idea of an algebraic theory of analysis than Grabiner seems to accept.

movement of a curve, and a solid body from the movement of a surface".<sup>26</sup>

For these continuous magnitudes Spehr stated the axiom (*Grundsatz*): "If a growing magnitude advances from one state to another... then it will pass through all intermediate states".<sup>27</sup> This is an axiom of applied science, and one would have expected Spehr to prove that all numerically continuous functions obey it. He failed to do this, although it could be easily done for the small class of functions he dealt with. Spehr seems to have considered this axiom a trivial property of his functions (which, in fact, it is). Other authors (Lagrange, M. Ohm) provided the necessary proofs.<sup>28</sup>

The applied part of Spehr's textbook contains two sections. The first one deals with the *general properties of fluents insofar they are derived from their known laws*. In the second section the laws of fluents are derived from their known properties. This section contains the applications of the general theory to geometry and mechanics. The purely verbal definitions of continuity, fluent etc. are used to interpret the equations, given in the arithmetical introduction (*analysis*), by means of examples from the applied disciplines of geometry and mechanics. Thus Spehr regarded continuity as a concept of the applied sciences, and not as a concept of the pure theory. Obviously, Spehr thought that this concept did not require a formal definition, and that he only had to make sure that the equations of theoretical analysis could be applied to geometric and mechanical phenomena; the verbal definitions served only as an aid for this work of interpretation.

In G. W. F. Hegel's *Science of Logic* (second edition of 1832) Spehr's book is mentioned in two passages, in which Hegel discusses the infinitesimal calculus. In one Hegel criticizes Spehr's definition of continuity as a mere "formal definition" (*formelle Definition*) which "expresses tautologically what the definitum is".<sup>29</sup>

Hegel is right if Spehr's definition is viewed as a conceptual foundation for a theory of continuity from which one could deduce theorems about continuous magnitudes. For this Spehr's definition is, of course, inadequate. It can only serve in a phenomenological way to provide some intuitive knowledge to discuss the question of which properties of the curves defined by the equations of pure analysis correspond to which phenomena of the world around us; it is a means of interpretation, similar to the Euclidean definitions of point, line and plane.

We thus arrive at the following general picture with regard to the problem

---

<sup>26</sup> Spehr [1826], 125.

<sup>27</sup> I.c., 127.

<sup>28</sup> Lagrange [1797], § 6. Ohm [1822], vol. 2, 123/4.

<sup>29</sup> Hegel [1832], 315.

of continuity. On the one hand, there is pure analysis as a universe of discrete symbolical forms (or, simply, formulae). If these forms are to be applied, a numerical interpretation becomes necessary. For this one must investigate the convergence of all infinite expressions. For the sake of numerical interpretation it must be shown that the functions given in the first step by symbolical expressions have adequate properties, i.e., properties which one would expect of geometric and/or mechanical objects. For example, it is expected that geometric curves are continuous, i.e., unbroken. For a given function  $f$  the difference  $f(x+h) - f(x)$  should become arbitrarily small for sufficiently small  $h$ . If a continuous function has a positive value at some point  $r$  and a negative value at another point  $s$ , then it should take on the value 0 for at least one point between  $r$  and  $s$  (intermediate value theorem). These properties are to be proved in the applied part where the numerical interpretation of the formulae is involved. Lagrange provided proofs for certain special cases.<sup>30</sup> In his textbook, M. Ohm provided proofs – conforming to modern standards – for analytic functions.<sup>31</sup> We repeat, that Spehr gave no formal proofs.

If formal series are accepted as basic objects of the pure theory, then even those proofs become correct in which the continuity of a function is derived by using its differentiability. Since functions defined by power series are differentiable *per definitionem*, using this property to prove their continuity is admissible. If the concept of a continuous function becomes the basic concept of the whole theory, and if functions are defined independently of any power series expansion (which is what Cauchy did in 1821), then the relation of both concepts becomes problematic. We know that it was not before the second half of the 19th century that mathematicians really understood that continuous functions are not in general differentiable. (For isolated points this was already obvious in the 18th century).

### 5. Theory and Applications Revisited

If we compare the approach of Lagrange, Spehr and Ohm, who did not use the concept of continuous function (though Lagrange and Ohm proved that their functions were continuous) with the approach of Cauchy who made continuous functions the basic objects of his theory, then we see that both lines of thought are equally legitimate from a logical viewpoint. The realm of the intuitively continuous is not mathematically encompassed by a formal

---

<sup>30</sup> The intermediate value theorem appears in Lagrange [1798], sections 2 and 6 as a method for solving equations  $f(x) = 0$ . Cf. Grabiner [1981], 69.

<sup>31</sup> Ohm [1822], vol. 2, 125/6, gives a proof of the intermediate value theorem which is fully acceptable in modern terms.



mathematical definition of continuity and Cauchy's theory with its formal concept of continuity was not *per se* nearer to applications. Rather, the applicability of such a definition to real phenomena remains a non-mathematical problem, to be solved by considerations coming from a particular area of applications. Hence the difference between a theory based on a formal concept of continuity and another one without such a concept is that both approaches result in different mathematical theories, and not that one theory is *a priori* more applicable than the other. Basing analysis on a formalized concept of continuity, is largely due to internal mathematical reasons and not to the notion that discrete mathematics is less applicable than continuous mathematics. Applications have an indirect and essentially mediating impact on the internal logic of mathematics. For example, they can produce a kind of pressure to include in pure mathematics new types of functions of hitherto marginal interest. A relevant example are functions definable by Fourier series. Such functions have played a considerable role in changing the foundations of analysis in the 19th century.

Many (but not all) 19th-century mathematicians, were aware that it would be an empiricist mistake to think that mathematical continuity would immediately encompass what we associate with continuity in nature. For example, in connection with his new construction of the real numbers  $\mathbb{R}$ , Dedekind wrote: "This property of the line [namely, that every cut corresponds to a point] is just an axiom by which we attribute continuity to the line, by which we imagine the line as continuous. If space has real existence, it need not necessarily be continuous; innumerable many of its properties would remain the same if it were discontinuous. And if we were sure that space is discontinuous, nothing would prevent us from making it continuous in our mind by filling its gaps if we wished to do so; ...".<sup>32</sup>

We conclude that Cauchy's theory was not superior to the older approach because it was better fitted, so to speak, to the nature of continuity, but – this is our thesis – because it provided essentially greater powers of representation and distinction. In the course of the 19th century functions not amenable to Lagrange's approach became representable and analyzable by Cauchy's methods. What favored the approach of Cauchy and Weierstrass was not the continuous functions but the construction of ever new classes of discontinuous functions, generated by more and more complicated means of representation. To (incorrectly) charge the older approach with circularity is to obscure the internal logic of mathematical development and to underestimate the true merits of Cauchy and of the newer analysis.

---

<sup>32</sup> Dedekind [1872], 11.

The elaboration of the algebraic approach to analysis by some members of the combinatorial school was quite modern from a philosophical viewpoint. The systematic distinction between theory and applications provided the opportunity for abandoning empiricist ideas about the nature of mathematical concepts. If we suppose that, in principle, continuity in the real world cannot be conceptually grasped but only approximated, then both approaches, Lagrange's and Cauchy's, have their respective advantages and disadvantages. Neither can claim an *a priori* higher epistemological status.

Every mathematical theory is a free creation of the intellect, and as such not a simple image of external phenomena. The German mathematician E. H. Dirksen defined analysis as a science having the peculiar property that "its objects as well as their determinations exist only insofar as they are produced by a free activity of the intellect; and this is the reason why in this field of knowledge nothing is recognized from the outside, but only from the way it is constructed".<sup>33</sup>

If mathematical concepts result from free constructions, then their application is a qualitatively new problem with its own demands and necessities. Although applications influence the development of theories, they must leave room for their internal development. Thus, scientists must consciously think about the connection as well as the difference between theory and applications. This situation yields a new understanding of the epistemological status of scientific theories that is clearly distinguishable from Kant's. While for Kant mathematics has an *a priori* sphere of meaning within the pure intuition of space and time, the new understanding of scientific theories admits concepts and theories whose sphere of meaning evolves in a complicated process of application.

This may have been Hegel's viewpoint when he tried to interpret the infinitesimal calculus within his *Science of Logic*.<sup>34</sup> He titled the relevant passage "The aim of the differential calculus derived from its application". Hegel distinguished clearly between an analysis of the internal relations of a theory and their justification through applications, and he emphasized this as a merit of Lagrange's approach. "... we owe to his [Lagrange's] method the essential insight that the two transitions necessary for the solution of the problem must be separated and each of the two sides must be treated and proved in itself".<sup>35</sup> At the end of the chapter he quoted with approval F. W. Spehr's views on the need to separate theory from applications.

---

<sup>33</sup> Dirksen [1845], III.

<sup>34</sup> cf. Wolff [1986], 224.

<sup>35</sup> Hegel [1832], 339.

With all necessary caution, I think that it is legitimate to describe the two conceptions of a theory discussed here in terms of modern philosophy of science. Kant's conception would correspond to what has been called the Statement View, whereas the ideas of Lagrange and the combinatorial school might be interpreted as representing the first step towards a Non-Statement View.<sup>36</sup> Historically, this was a necessary intellectual condition for the development of pure mathematics during the 19th century.

### References

- Becker, O.; [1927]: "Mathematische Existenz. Untersuchungen zur Logik und Ontologie mathematischer Phänomene". *Jahrbuch für Philosophie und phänomenologische Forschung* 8. 2nd. edition, Tübingen: Niemeyer 1973.
- Becker, O.; [1975]: *Grundlagen der Mathematik in geschichtlicher Entwicklung*. Frankfurt.
- Bekemeier, B.; [1987]: *Martin Ohm (1792–1872): Universitätsmathematik und Schulmathematik in der neuhumanistischen Bildungsreform*. Studien zur Wissenschafts-, Sozial- und Bildungsgeschichte 4. Göttingen.
- Dedekind, R.; [1872]: *Stetigkeit und irrationale Zahlen*. Braunschweig. Reprint of the 7th edition, Braunschweig 1969.
- Dirksen, E.H.; [1845]: *Organon der gesamten transcendenten Analysis*. 1.Theil. *Transcendente Elementarlehre*. Berlin.
- Euler, L.; [1748]: *Introductio in analysin infinitorum. Tomus primus*. Lausanne. In: *Leonhardi Euleri Opera Omnia* I<sub>3</sub>.
- Euler, L.; [1755]: *Institutiones calculi differentialis*. Basel. In: *Leonardi Euleri Opera* I<sub>10</sub>.
- Grabiner, J. V.; [1981]: *The Origins of Cauchy's Rigorous Calculus*. Cambridge, Mass.
- Gudermann, C.; [1825]: *Allgemeiner Beweis des polynomischen Lehrsatzes ohne die Voraussetzung des binomischen und ohne die Hilfe der höheren Rechnung*. Schulprogramm. Cleve.
- Hankel, H.; [1871]: "Grenze". In: *All. Encykl. der Wiss. und Künste*, erste Sect., neunzigster Theil. Leipzig, 185–211.
- Hegel, G. W. F.; [1832]: *Wissenschaft der Logik*. 2.edition. G. W. F. Hegel: Werke 5. Frankfurt-M.: Suhrkamp.
- Herder, J. G.; [1799]: *Vernunft und Sprache. Eine Metakritik zur Kritik der reinen Vernunft. Erster Teil*. In: J. G. Herder, *Sämmtliche Werke*, ed. by B. Suphan, vol. 21. Berlin 1881, 1–190.
- Hindenburg, C. F. (ed.); [1796]: *Der polynomische Lehrsatz, das wichtigste Theorem der ganzen Analysis: nebst einigen verwandten und anderen Sätzen, neu bearbeitet und dargestellt v. Tetens, ... Zum Druck befoerdert und mit Anmerkungen, auch einem kurzen Abrisse d. combinatorischen Methode und ihrer Anwendung auf die Analysis versehen*. Leipzig.
- Jahnke, H. N.; [1990]: *Mathematik und Bildung in der Humboldtschen Reform*.

---

<sup>36</sup> cf. for this distinction: Sneed [1971], Stegmüller [1973].

- Studien zur Wissenschafts-, Sozial- und Bildungsgeschichte der Mathematik vol. 8. Göttingen: Vandenhoeck & Ruprecht.
- Kant, I.; [1781]: *Kritik der reinen Vernunft*. Riga. In: W. Weischedel (Hrsg.), *Kant Werke II*, Wiesbaden 1956.
- Lagrange, J. L.; [1770]: "Nouvelle méthode pour résoudre les équations littérales par le moyen des séries". *Mémoires de l'Académie royale des Sciences et Belles-Lettres de Berlin XXIV*. In: *Oeuvres de Lagrange III*, 5–73.
- Lagrange, J. L.; [1797]: *Théorie des fonctions analytiques*. Paris. In: *Oeuvres de Lagrange IX*.
- Lagrange, J. L.; [1798]: *De la résolution des équations numériques de tous les degrés*. Paris. In: *Oeuvres de Lagrange VIII*.
- Newton, I.; [1676]: Letter to Oldenburg 24, October, 1676. In: H. W. Turnbull (ed.), *The Correspondence of Isaac Newton*, vol. II: 1676–1687. Cambridge 1960, 110–161.
- Ohm, M.; [1822]: *Versuch eines vollkommen consequenten Systems der Mathematik, erster und zweiter Theil*. Also as: *Lehrbuch der Arithmetik, Algebra und Analysis. Nach eigenen Principien*. Berlin.
- Pfaff, J. F.; [1795a]: "Analysis einer wichtigen Aufgabe des Herrn de la Grange". *Archiv der reinen und angewandten Mathematik I*, (1. H.), 81–84.
- Pfaff, J. F.; [1795b]: "Ableitung der Localformel für die Reversion der Reihen, aus dem Satze des Herrn de la Grange". *Archiv der reinen und angewandten Mathematik I* (1. H.), 85–88.
- Rothe, H. A.; [1793]: *Formulae de serierum reversione demonstratio universalis, signis localibus, combinatorio - analyticorum vicariis exhibita*. Leipzig.
- Rothe, H. A.; [1795]: "Lokal- und combinatorisch-analytische Formeln für höhere Differenziale". *Archiv der reinen und angewandten Mathematik I* (4. H.), 431–449.
- Sneed, J. D.; [1971]: *The logical structure of mathematical physics*, Dordrecht.
- Spehr, F. W.; [1824]: *Vollständiger Lehrbegriff der reinen Combinationslehre mit Anwendungen derselben auf Analysis und Wahrscheinlichkeitsrechnung*. Braunschweig.
- Spehr, F. W.; [1826]: *Neue Principien des Fluentencalculs, enthaltend die Grundsätze der Differential- und Variationsrechnung unabhängig von der gewöhnlichen Fluxionsmethode, von den Begriffen des unendlich Kleinen oder der verschwindenden Größen, von der Methode der Grenzen und der Functionenlehre, zugleich als Lehrbuch dieser Wissenschaft dargestellt, und mit Anwendungen auf analytische Geometrie und höhere Mechanik verbunden*. Braunschweig.
- Stegmüller, W.; [1973]: *Theorienstrukturen und Theoriendynamik. Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie*, vol. II, 2. Halbband. Berlin-Heidelberg-New York.
- Wolff, M.; [1986]: "Hegel und Cauchy. Eine Untersuchung zur Philosophie und Geschichte der Mathematik". In: Horstmann, R.-P. & Petry, M. J. (eds.), *Hegels Philosophie der Natur*. Stuttgart: Klett-Cotta.

## Constructivism and Objects of Mathematical Theory

MICHAEL OTTE (Bielefeld)

Most philosophers of mathematics have regarded both human history in general and the history of mathematics in particular as epistemologically irrelevant. Mathematics seems to be an intellectual field in which historical development is swallowed up by the latest state of the art, at the same time preserving what remains worthwhile. Many believe that "telling the story of any theoretical subject  $x$  is not a conceptually distinct undertaking from describing the theory of  $x$  ... Worse yet, mathematics ... does not admit a history in the same sense as philosophy or literature do" (Rota et al. 1986, 157).

If, however, one is to be able to actively develop or use any mathematics in a meaningful way, one has to place it into one's own context. For the majority of people, this implies an endeavour to find connections between mathematics and other fields of experience, and therefrom results an interest in history. I believe, in fact, that the ideas we carry with us about what human history is will influence our conceptions concerning mathematical epistemology.

Mathematics and logics did not just emerge from a meta-analysis of societal exchange, of communication and language, as the logical empiricists seem to believe – a belief that leads them to maintain an absolute distinction between the analytic and synthetic and to take the laws of logic and the propositions of pure mathematics to be analytic. There exist in fact two alternative comprehension schemata, "which dominate contemporary philosophical culture: the paradigm of language and the paradigm of production" (Markus 1986). Since the early 19th century, there have existed two ways of thinking in mathematics that more or less correspond to these schemata. These two ways of thinking manifested themselves, for instance, in the criticism of Kant by Bolzano on the one hand and by Hegel and Grassmann on the other.

### *1. Apriorism versus Empiricism in Mathematics*

The rationalism of the 17th and 18th centuries was rooted in God: "Every rea-

lity must be based upon something existent", said Leibniz, "if there were no God there would be no objects of geometry" (quoted from Lovejoy 1936, 147). It appears quite natural to the contemplative mind to search for the "existent" (i.e., that being which is no longer reduced to something else but speaks for itself) in the spiritual realm, and to lay a base for true knowledge there. "If someone should reduce Plato to a system, he would render a great service to the human race", said Leibniz, "and it will be seen that I have made some slight approximation to this". In addition to this, the thinking of this period "is determined by what is thought, by its object" (Gaidenko 1981, 57); it is ontological thinking. The ontological character of knowledge is responsible for the fact that mathematics during this period was synthetic in character, contrary to being permanently concerned with "analysis" (Boutroux 1920). In 18th century usage, 'the synthetic' and 'synthesis' were related both to the logical and to the empirical, whereas mathematics was regarded as an "analytical language" (Condillac) and as a calculus, that permits operations with objects that do not "really exist" (L. Carnot). But this was at times considered with some uneasiness. So, even though mathematics in the Modern Age was referred to as analytical since Descartes (particularly in its orientation towards the generalizability of the mathematical method, which was possible due to calculus), one has to interpret this in the context of the philosophical ontologism and representationalism of the Classical Age (Bos 1984). By placing the emphasis more on the constructiveness of mathematics than on any other aspect, Kant had emphasized the problem that was central to the 19th century: the relationship between empiricism and apriorism.

To goal-driven activity, only the empirical, as well as that which has a clear empirical purpose appears as really existent and indubitable. Knowledge is meant to have a function and is not autotelic. In order to understand the power and efficacy of mathematics, Kant intended to reconcile empiricism and apriorism, and therefore he searched for the regulatory principles of cognitive activity. Therefrom the forms of pure intuition conceived of in terms of space and time. On the one hand, Kant was the first philosopher to concentrate on activity or construction as the one fundamental element of epistemology. It is by means of his own activity that man establishes relations with the social and objective world. On the other hand, activity or construction conceived of as a mere empirical process and the functionalism associated with such a conception do not provide foundations for any idea of truth (Kant highly appreciated Hume's reflections with respect to this problem). Space and time as forms of pure intuition were supposed to provide foundations that were at the same time compatible with the insight that cognition has an essentially active character and that knowledge is based on construction.

As early as 1810, Bolzano criticized Kant's elucidation of the arithmetical proposition " $7+5=12$ ". In an appendix to his *Beiträge zu einer begründeteren Darstellung der Mathematik* (1810) B. Bolzano, while replacing the original theorem by the simpler one of " $7+2=9$ " for the sake of convenience, comments on Kant as follows: "Most of the theorems of arithmetics are, as Kant correctly remarked, synthetical theorems. But who does not feel how forced was what Kant had to proclaim in order to give his theory of intuition general validity: that these theorems are also based on an intuition, and (how could it be otherwise) on the intuition of time?"

Proving the theorem  $7+2=9$  "shows no difficulty if one assumes the general theorem that  $a+(b+c)=(a+b)+c$ , that is, that in an arithmetical sum, one regards only the quantity but not the order (certainly a concept different from that of temporal sequence) of the items. This theorem even excludes the concept of time instead of assuming it. If we accept it, however, the above can be proved as follows: that  $1+1=2$ ,  $7+1=8$ ,  $8+1=9$  are mere explanations and arbitrary theorems. Thus  $7+2=7+(1+1)$ ,  $=(7+1)+1=8+1=9$ ". So far Bolzano's reasoning by recurrence. Evident here is already the transition from the arithmetic functions as mere processes and activities to their transformation into particular objects of consideration. This process is continued in Hermann Grassmann's 1861 arithmetics textbook, in which the recursive justification of arithmetic functions is used for an axiomatic foundation of arithmetics (Otte 1990). The overall intention of Bolzano's critique of Kant is directed towards the purport of an alleged empiricism. This becomes even clearer with respect to the intuition of space. Kantian philosophy takes, as Bolzano says, "those intuitions which shall be a quite particular addition to mathematical definitions of concepts to be nothing but an object subordinated to the definition of a concept in geometry, an object, which our productive imagination is to add to the definition provided..."

It has to be said that "what is demanded here may well apply to many but by no means all concepts pertaining to geometry. So, for example, the concept of an infinite line is also a geometric concept, which therefore also has to be explained geometrically. And nevertheless the productive imagination can certainly not create an object which corresponds to this concept. For we cannot draw an infinite line by means of any imagination, but we can and we have to think it by means of reason only" (B. Bolzano 1975, 76/77). This implies that geometry is, contrary to Kant's proposition, also analytic.

Hegel, in contrast, holds Kant to be guilty of a certain subjectivism. Unlike Bolzano or the Platonist view of mathematics in general, Hegel accepts Kant's argument that the activity of the subject plays an essential

role in the process of knowledge, and that mathematics in particular must be defined in genetic terms – must be described with respect to its active development. Hegel, however, criticizes Kant from the perspective of absolute consciousness. For Kant, the strength of mathematics lies in the very fact that mathematics to a certain degree represents an ideal of cognition in which understanding is based on synthesis and construction and on direct intuitive evidence to which the construction leads. In this sense, mathematics represents a direct, quasi local knowledge, that is a knowledge whose reasons are not to be looked for beyond the context of what is directly accessible.

Mathematics seems to be a very simple, direct knowledge. A mathematical text says everything it has to say and it may be understood literally. Mathematics is direct in the sense of the "minimal loop". If I say "p", this means "p". The referential meaning is reduced to the predication " $a = a$ ". Everything means itself and does not refer to some other. This directness, of course, is also some kind of *rigidity*. In its history, mathematics has bound itself primarily to an ideal of cognition that equates cognition either with seeing or with logics.

In this way, however, as Hegel believes, mathematics becomes a merely subjective knowledge, and Kant's constructivism itself remains within the subjective – an attitude already expressed in the very fact of Kant's estimation of mathematics. Hegel says: "Thoughts, according to Kant, although universal and necessary categories, are *only our* thoughts separated by an impassable gulf from the thing, as it exists apart from our knowledge. But the true objectivity of thinking means that the thoughts, far from being merely ours, must at the same time be the real essence of the things and of whatever is an object for us" (Hegel 1832, § 41).

Recently, Kitcher has attacked Kant's constructivism using arguments similar to Hegel, claiming that mathematics, according to Kant, simply describes properties of "transient and private mental entities" (Kitcher 1984, 55). Kant may perhaps be accused of psychologism, being interested in mathematical truths because they necessarily appear to be true and are not just true. On the other hand Kant has repeatedly stressed that our consciousness of self, our "internal sense", is never an immediate experience but must be mediated by external objects or means of cognition. The general conditions of construction that Kant addresses obviously depend on an overall system of culturally produced means of representation and knowledge – of which our senses only make up a small part. In this perspective, it seems doubtful whether there exists a Kantian thesis according to which "our psychological constitution dictates the geometrical structure of experience" (Kitcher 1984, 55). This



geometrical structure is in fact determined by the prerequisites of measurement. From the latter it follows that even in Einstein's relativity theory a locally Euclidean structure of space is presupposed (Borzeskowski / Wahsner 1979, Weyl 1968, vol 2, 40).

Taking all this together, one is once more inclined to follow Kant in his endeavours to establish activity as connection between empirical and psychological reality. The essential problem coming to the fore at this point consists in the conceptualisation of the system of means of mathematical activity as the structure of that activity itself.

Reference to the means of cognition makes it possible to conceive of the limiting conditions and regulatory principles that Kant sought in the pure forms of intuition in an evolutionary instead of a nonhistorical manner. All modern philosophy since Kant has found itself confronted in some way or other with the fact that "the regulatory principles deriving from philosophy are considered at one point as the product of the evolution of cognition, and at another as its indispensable condition" (Amsterdamski 1975, 175).

## 2. The Role of Metatheoretical Principles

If we consider the hierarchy: "meta-theory – theory – external reality", it seems that for 17th and 18th century thought the first relation was well established, whereas the second remained strangely unrelated. During the 19th century the situation seems reversed. In Grassmann's hands, for instance, *axiomatics* turns into a system of meta-theoretical statements in the service of mathematics conceived of as a science of forms.

Axioms for Euclid are immediately comprehensible, content-related fundamentals of theory from which the theory can be more or less logically deduced. That is to say, axioms justify themselves by reference to the objective fundamentals of the theory. This ontologic foundation of knowledge which prevailed well into the 19th century had the effect that different or alternative theories of a subject matter could not be conceived of.

In Grassmann's *Ausdehnungslehre* of 1844, axiomatics is a system of transcendental requirements for the development of every individual mathematical theory. This new role of axiomatics is the same as that of the symmetry principles and conservation laws in physics. As Wigner emphasized in his Nobel Prize speech in 1963, we have also in physical knowledge a layering – and the laws of nature have the same function in this hierarchy with respect to events as the symmetry principles have with respect to the laws of nature. This functionality reveals itself, above all, in dynamism. In other words, if we knew all the laws of nature, then the invariance properties of these laws would

not furnish new information, and the symmetry principles would only contain a more or less superfluous postclassification of the same. The same is true for the relationship between natural law and real event. If we knew all the facts, the natural laws would be more or less a superfluous description.

An analogous message may be derived from the so-called "theoretician's dilemma" (cf. Tuomela 1983, 6) which maintains that theoretical concepts are superfluous: Theories have a right to exist only insofar as they have transcendental character with respect to the empirically perceptible, just as axiomatic meta-principles have only a right to exist on account of their transcendental character with respect to the respective theories in question. And this "right to exist" only shows itself in the dynamics of cognition.

The essential question then is what role the objects under study play in guiding the dynamics of theorizing. Grassmann's answer was that this dynamics is determined by an interaction of the "symmetry principles" of his general theory of forms and an intuitive idea or preconception of the object field under study. The conception of mathematics as a science of forms depends on the availability of different interpretations or intended applications.

The legitimating meta-level discourse evolves alongside, or is based on, the question as to what essentially constitutes man's relations to external reality or what makes up his being within the world. The prevalent answers, although differing in conceptualisation and in detail, repeatedly refer to ideas like "*activity*", *social practice*, *construction*, and so forth. I want to sketch this in the following two examples: in the example of the concept of *function*, and in the problem of *equality*. The function concept was fundamental to the emerging conceptual approach in mathematics, as were formal identities to the constructivist approach.

J. T. Merz, in his monumental *History of European Thought in the 19th century*, has written: "The conception of correspondence plays a great part in modern mathematics. It is the fundamental notion in the science of order as distinguished from the science of magnitude. If the older mathematics were mostly dominated by the needs of mensuration, modern mathematics are dominated by the conception of order and arrangement" (Merz 1903, 736).

In the light of the differences between the conceptual versus the constructivist approaches to mathematics, I may perhaps add to this a qualification, claiming that "the needs of mensuration" in fact provided a very strong stimulus for the evolution of modern axiomatics that originates with the work of Grassmann. Grassmann was interested in conceptions of structure and arrangement because he wanted to find out in which way an object field has to be conceptualized such that processes of measurement can be applied at all.

### 3. The Concept of Equality

On the basis of Leibniz's principle of identity of indiscernables, the *equality* of two objects is determined by the fact of them having all features in common, or, in other words, that they produce the same values as arguments for all functions.

$$x = y \text{ iff } f(x) = f(y) \text{ for all } f$$

Equality is an equivalence relation that must be compatible with all functions and is determined by this compatibility. Such a requirement is not operative, and in fact Leibniz's principle contains a succinct expression of classical ontologism. The ultimate goal, which is, in general, only to be accomplished by God through an infinite analysis, lies in the determination of the individual substances.

The constructivism of modernity is in contrast a sort of relativism. According to this latter view, it does not seem necessary for us to identify the individual objects in every respect. Numbers, for instance, as Niiniluoto has said "can be well-defined relative to their relational arithmetic properties but indefinite relative to other properties", (Niiniluoto 1992, 65). In a constructive view, it is the identification of the objects, not the difference between them, that must be specified. Every such identification will be relative and perspective-dependent. Cassirer, in his work *Substanzbegriff und Funktionsbegriff*, wrote in the same vein:

"While the empiristic doctrine regards the 'equality' of certain contents of presentation as a self-evident psychological fact which it applies in explaining the formation of concepts, it is justly pointed out in opposition that the equality of certain elements can only be spoken of significantly when a certain 'point of view' has been established from which the elements can be designated as like or unlike. This identity of reference, of point of view, under which the comparison takes place, is, however, something distinctive and new as regard the compared contents themselves. The difference between these contents, on the one hand, and the conceptual 'species', on the other, by which we unify them, is an irreducible fact; it is categorical ...". (Cassirer 1910, 33).

Cassirer derives from this a difference in principle between objects and concepts, or, in other words, he interprets Kant's insight that concepts are not just simply acquired from objects in a process of abstraction by emphasizing a difference in principle between theories and the world of objects to which they refer. If, however, this difference is taken to be absolute in the sense of a full freedom for mathematical construction, and mathematical concepts are no longer co-determined by extension, then the problem of equality takes on a merely

formal character, or disappears altogether. If it is assumed, on the other hand, that an axiomatic theory cannot be seen as fully intensional, but that it refers to one of its various objective contents, then specific functions in the sense of the above compatibility requirement are selected for this theory. The equality relation, then, need only be compatible with particular theory-characteristic functions. One could even say it need possibly only be compatible with one function. In the economic equation, "1 suit=2 pairs of shoes", the suit and the pair of shoes have only their economic value, in common and nothing else.

Thus the requirements of mathematical construction are not arbitrary. That is to say, the range of functions that distinguish the identity relation are not arbitrarily selectable, as long as the theory is supposed to be applicable. I have, shown, for example, in examining the different approaches of Grassmann and Leibniz to the problem of "geometrical characteristics", that the congruence relation chosen by Leibniz according to an Euclidean tradition did not function as geometrical equality because it is not compatible with certain requirements of measurement, namely that it does not give rise to extensive geometrical quantity in the sense outlined by Lawvere (Lawvere 1992). It was exactly this fact that led Grassmann to choose to area equivalence (in the two-dimensional case) as an identity-constituting equivalence relation (Otte 1989, 24/25). Congruence is to be distinguished from geometrical equality already on Euclid's axiom: "If equals be added to equals, the wholes are equal" (Euclid's *Elements*. Book I). This axiom is true with respect area equivalence, whereas in the case of congruence the "whole" depends on how it is constituted from parts. Grassmann models the conception of space by means of a vector space of arbitrary finite dimension, endowed with a determinant function  $D$ . Such a space  $(V, D)$  is sometimes called a Peano space (Rota et al. 1985). A Peano space can be viewed geometrically as a vector space in which an oriented volume element is specified.

Space, for Leibniz, is a relative ordering of objects. But for the 19th century, every such ordering must be compatible with the activity of measurement and with the structure of this activity.

Every theory constitutes a particular context of mathematical knowledge, and we have something similar to a context-dependent conception of mathematical meaning that is only modified, changed or developed so that this theory may be used in new object-worlds. Equality relations are equivalence relations that are compatible with certain functions or operative structures such that they can be employed by a process of "definition through abstraction" to construct the basic ontology of the axiomatized theory in question. Every theory has its particular ontology, about which it speaks. The term ontology no longer designates that which exists as such in the world, but refers to those

aspects of reality about which we have the intention as well as the means to speak in a mathematically meaningful way. By the term ontology, the set of those entities is described whose existence is stipulated by the theory.

As the question as to whether two theories speak about the same object-field is usually settled by the course of goal-oriented social practice, it seems doubtful whether "an adequate understanding of the problem of mathematical identity requires a new and still missing formal theory that will describe the mutual relationships that obtain between formal systems that describe the same object" (Rota et al. 1988, 377).

In general, how a person conceives of the objects of a theory certainly depends on the purposes and goals that they might bring to it. For instance, from the perspective of the user (where use also includes theoretical use), the theories are strictly differentiated from the object areas to which they refer, because the user sees the object in his own perspective. Whereas, from the perspective of the theory designer, from the perspective of the activity itself, the objects of the theory seem to be identical with the way in which they are present within the theory. Objectivity is based on social practice in as much as it presupposes alternative approaches to an object.

Since Bolzano and Frege, an equation  $a=b$  is usually interpreted as referring to the same thing in two different ways.  $a$  and  $b$  differ intensionally but are identical extensionally. Such an interpretation taken absolutely may forget that, within the dynamics of the process of cognitive activity, it is generally the intensions that count and not the extensions. For to be able or unable to solve a problem, it is, for instance, of the greatest importance how that problem may be presented. The same is true of any knowledge that we want to apply.

The view presented here has the advantage that it becomes clear, for example, that consistency proofs need only be of a contextual nature for mathematicians; that the mathematician is not interested in global consistency proofs. Only when one wants to specify the transcendental requirements of mathematical theory-formation absolutely does the problem of *global* consistency proofs appear. In this sense, Gödel showed that there were no absolute transcendental requirements for mathematical construction. In the light of the argument above, this means that mathematical thought is objectively oriented and objectively determined in its development, and that its diversity is, among others, attributable to the complexity of the world.

#### 4. The Concept of Function

Set-theoretical mathematics does not consider identity of individual objects but

only equality of functions, predicates, sets, and so forth; that means, higher-order entities on the basis of the axiom of extensionality. The identity of two functions is established in this way, that is, two functions are identical if they produce the same values for the same arguments:

$$f = g \text{ iff } f(x) = g(x) \text{ for all } x$$

(according to type theory,  $X_{k+1}(X_k)$  would replace  $f(x)$ ; i.e.,  $X_{k+1}$  contains  $X_{k+1}$ ; (cf. E. Beth, 1959, 226). The identity of the elements on the fundamental level is assumed *per se*.

On the other hand, the basis of Leibniz's principle of identity of indiscernables could be used to establish the identity of the arguments by the fact that they have all their features in common, or, in other words, that they give the same values as arguments for all functions:

$$x = y \text{ iff } f(x) = f(y) \text{ for all } f$$

A dissymmetry is obvious which shows that individual objects are more abstract than functions and other entities of higher logical type. The dynamics of constructive mathematical activity removes this dissymmetry by generally using higher-order types to establish the identity of lower types, and in this manner explains the abstract by means of the less abstract, although this simultaneously implies that very often the phenomenologically less well known is employed to explain the better known, for instance, Newton's laws explaining the motion of bodies or Ohm's law to explain electrical phenomena. "True, in the past, when people tried to explain phenomena animistically they did not use general laws but 'explained' unknown phenomena by known (or seemingly known) ones. The situation was similar with the mechanistic explanation, e.g. the consideration of an organism as a machine" (Krajewski 1977, 30).

To make such an analytic approach feasible, we use a lot of metaphorical terminology. When one speaks, for example, about electrical current, one does not confuse electricity with hydrodynamics. What appears abstract to the constructive or the contemplative mind respectively may be two very different matters.

With respect to the function concept itself, we not only define the equality of functions by means of the extensionality axiom but also use the other constructive way of employing certain functionals, "functions of functions", to establish it. The transition of mathematics from the 18th to the 19th century is characterized in general by the fact that the objects were no longer given first. For example, given a linear function (represented, say, by the symbolic expression  $f(x)=ax$ ), one tried in the old synthetic approach to read

off features of the presented object (e.g., the functional equation  $f(x+y)=f(x) + f(y)$ ). Or, to take the example of group theory: The concept of group first meant some sort of transformation group and the elements were considered as particulars. Only later did the abstract concept of group arise. Starting from this, propositions stating, for example, that a discrete or continuous group with such and such properties has an isomorphic linear representation constituted an important part of group theory and its applications. The 19th century saw the beginning of an approach that defined the objects of mathematical activity in terms of functionality (e.g. the above functional equation and the continuity of the function) and constructs representations or other features of the object from those given (e.g. the representation  $f(x)=ax$ ).

During the 17th and 18th centuries, the concern was about objects, whereas one was completely free with respect to methods. In the 19th century, the situation was reversed. Everything now seemed able of being treated mathematically, at least in principle, whereas standards of method became more and more specified. The rationality of science became the functionality of its methods. Specialized pure mathematics was increasingly based on proof analysis and became analytic. In an analytic approach, the objects of a theory are identical with their definitions as provided by the axiomatic foundation of the theory. The synthetic approach prevailing during the 17th and 18th centuries (Boutroux 1920), despite being called "analytic", relied on the complete set of properties of the object, which it did not and sometimes certainly could not explicitly describe. As has been said: The linear function or the "straight line" and so forth are given by  $y=ax$ ;  $y=ax+b$ , and so forth.

(A very instructive modern example showing the interaction between the analytic and synthetic is provided by group theory: On the one hand, one has groups being given an axiomatic description, and, on the other, one employs, for instance, linear representations of them. A model like a linear representation adds to our information in as much as the *individual* elements of the group are provided with additional properties that were not present in the abstract definition of the group and that can now be used by the mathematician. In this manner, group theory becomes analytical as well as synthetical).

It is in such a perspective that Cauchy's achievement (from whose *Course d'Analyse* the above example has been taken) stands out clearly. "A melange of methods based on limits, power series or differentials together with a ... combination thereof was replaced by a single doctrine founded on a univariate theory of limits ...". (Grattan-Guinness 1970, 1286). This new concentration on "method" liberated method and at the same time caused a gradual neglect of the indispensability of the continuous or of a continuity principle for adding meaning to the syntactical and operative – be it in the sense of intended

applications or as an intuited context. In 1895, Felix Klein described the development of an arithmetization of mathematics, which was an expression of this concentration on method, as follows: "Proceeding from the observation of nature and with the goal of explaining it, this spirit from which modern mathematics was born has concentrated on a philosophical principle, the principle of continuity. This is the case for the great pioneers, for Leibniz, Newton; this is the case throughout the 18th century that was truly a century of discovery for the development of mathematics. ... In Gauß the perception of space, in particular the perception of the continuity of space, is still used as an unquestioned proof. A closer study demonstrated not only that this was based on much that was not proven but also that the perception of space had too hastily led to theorems being considered to have a general validity they do not process. From this follows the demand for exclusively arithmetized proofs. Only that which can completely be proved to be identical by means of arithmetical calculation should be viewed as the property of science".<sup>1</sup> (Klein 1928, 143/144)

Klein's statement points to the essential role of the "principle of continuity", which was a fundamental principle of the 18th and 19th century thought (cf. Lovejoy 1932), in the development of modern mathematics. This principle was an example of a conception of a context for the development of mathematical construction. This is the reason for the relation observed by Bochner between the function concept and the continuity concept. "...It is a most significant fact that, in relatively recent Western Thought, the conceptions of function and of continuity have evolved simultaneously and in close intellectual interpenetration with and dependance on each other, in fact or perhaps only in intent" (Bochner 1974, 845).

During the transition from the 18th to the 19th century, the principle of

---

1 Von der Naturbeobachtung ausgehend, auf Naturerklärung gerichtet, hat der Geist, aus dem die moderne Mathematik geboren wurde, ein philosophisches Prinzip, das Prinzip der Stetigkeit an die Spitze gestellt. So ist es bei den großen Bahnbrechern, bei Newton und Leibniz, so ist es das ganze 18. Jahrhundert hindurch, welches für die Entwicklung der Mathematik recht eigentlich ein Jahrhundert der Entdeckungen gewesen ist. ... Bei Gauß wird die Raumschauung, insbesondere die Anschauung von der Stetigkeit des Raumes noch unbedenklich als Beweisgrund benutzt. Da zeigte die nähere Untersuchung, daß hierbei nicht nur vieles Unbewiesene unterlief, sondern daß die Raumschauung dazu geführt hatte, in übereilter Weise Sätze als allgemeingültig anzusehen, die es nicht sind. Daher die Forderung ausschließlich arithmetischer Beweisführung. Als Besitzstand der Wissenschaft soll nur angesehen werden, was durch Anwendung der gewöhnlichen Rechnungsoperationen als identisch richtig klar erwiesen werden kann.



continuity changes from a metaphysical to an epistemological status. It no longer represents a characteristic of matter as such but of matter as it is presented in our thought. It represents the continuity of all possible representations, or of all possible perspectives on an object. This idea is manifest in the writings of Carnot and Poncelet, among others, and it is fundamental to the philosophy of Peirce. According to him, the great characteristic of nature is diversity and arbitrary heterogeneity. Mental action, on the other hand, is characterized by a tendency to generalization, and generalization always employs some principle of continuity (Peirce 1965, 6.101).

But there is more to this principle. Set-theoretical mathematics is based on a general statement of existence. Constructivism denies the possibility of such a statement in as much as all our knowledge is relative to the nature of the human mind. Peirce, for instance, says: "What I propose to do ... is, following the lead of those mathematicians who question whether the sum of the three angles of a triangle is exactly equal to two right angles, to call in question the perfect accuracy of the fundamental axiom of logic.

This axiom is that *real things exist* or in other words, what comes to the same thing, that every intelligible question whatever is susceptible in its own nature of receiving a definitive and satisfactory answer, if it be sufficiently investigated by observation and reasoning. This is the way I should put it; different logicians would state the axiom differently. Mill, for instance, throws it into the form: *Nature is uniform*" (Peirce 1986, 545 f.).

We may observe at this point that the principle of continuity served a similar purpose to that of Platonism in pure mathematics, namely, to be assured that there is something to be known, that the world can be known. This purpose or function may be achieved, however, without necessarily retaining the ontological interpretation of the principle of continuity that had been so prominent during the 17th and 18th centuries.

The ontological interpretation of the principle of continuity was connected to the static worldview of the Classical Age (Lovejoy 1936) and its importance decreased as that of the conception of evolution grew. An evolutionist perspective introduces an element of indeterminacy or absolute chance in nature. It makes no sense to conceive of, for instance, the mathematical function-concept as of a direct image of a supposed causality or regularity of nature, as was the case during the 18th century. In this respect, Lagrange and Cauchy did not speak the same language. For instance "no matter how much Lagrange may assert and insist that a function is for him an 'abstract' mathematical object, in his thought patterns it somehow is residually a mechanical orbit or perhaps a physical function of state; whereas in Cauchy,

orbits and forces and pressures are always functions, as they are for us today" (Bochner 1974, 837).

However, just as it is unreasonable to reify the propositions of theory, it is absurd to conceive of their objectivity as an outgrowth of mere formal logic. Any evolutionary or historical perspective presupposes the Kantian project of mediating between empiricism and rationalism or apriorism. Absolute necessity and absolute chance are indistinguishable anyway (cf. Laplace, *Philosophical Essay on Probability*); and this shows that the means available to the cognitive activity delimit the space of epistemological possibilities at hand. The means of cognition, however, evolve alongside its objects, because the proof of the pudding is in the eating.

The new mathematically abstract view of the function concept was, as said already, inseparably bound together with the "continuity principle" (cf. Leibniz 1966, 84 ff. and the comments of his editor Ernst Cassirer). This principle introduces certain descriptive assumptions into the functional relation. A functional relation is continuous if a "small" variation in input causes a correspondingly restricted variation in output. In particular, the determinism arising from this binds the concept of the continuous function closely to the concept of law in classical natural science (cf. J. Gleick 1987, for the limits of this determinism).

On the one hand, the concept of function is radically operationalized and viewed as a "black box" that transforms "inputs" into "outputs". On the other hand, this radical operationalization takes place according to the requirement that reality is not chaotic but structured according to laws. This in turn seems to belong to the requirements for the application of mathematics to reality, which is that mathematics should function within the framework of the scientific knowledge of the world.

For Leibniz, this law of continuity was fundamental directly on account of the reason presented here, namely that it gives expression to a prerequisite for the applicability of mathematics, or the acquisition of knowledge about reality. This requirement is that reality be structured according to laws. If one denied the continuity principle, says Leibniz, "the world would contain hiatuses, which would overthrow the great principle of sufficient reason, and compel us to have recourse to miracles or pure chance in the explanations of phenomena" (quoted from Lovejoy, p. 181).

The role of the continuity principle in the formation of the function concept is particularly revealed in the fact that it was only a sufficiently abstract and general view of mathematical functions that made it possible to derive an interaction between the complementary aspects of the function as operation or rule on the one hand, and as a given causal relationship on the other. As late

as 1748, Euler defined a function in his *Introductio in analysin infinitorum* as an "analytic expression which is made up somehow of variable and constant number quantities", and believed that continuous functions are exactly those that may be represented in closed form by such an analytical expression.

This concept realism that transforms the fundamental feature of continuity into a feature of the symbolic "manifestation" of the function, thus considering it not as constitutive for the function itself, leads to great difficulties and inconsistencies, as the same function could simultaneously be labelled both continuous and discontinuous. "In the works of Euler and Lagrange" wrote Cauchy, "a function is called *continuous or discontinuous*, according as the diverse values of that function, ... are or are not produced by one and the same equation.... Nevertheless the definition that we have just recalled is far from offering mathematical precision; for the analytical laws to which functions can be subjected are generally expressed by algebraic or transcendental formulae ... and it can happen that various formulae represent, for certain values of variable  $x$ , the same function: then, for other values  $x$ , different functions" (quoted from Grattan-Guinness 1970, 50 ff.).

Certain fundamental features, such as continuity, could therefore only be assigned to a more abstract concept of a functional relation, a concept that must be acquired in a process of definition, by abstraction from equivalence classes of symbolic representations.

This means that operativity or functionality itself must be comprehended, as must be the conception of correspondence. This relativates the connection between the function concept and its symbolic representation. Lobachevski (1793–1856), for instance, wrote in 1834:

"The general concept requires that the function of  $x$  be called a number, which is given for all  $x$ , and changes progressively with  $x$ . The value of the function can be given either by an analytical expression, or by a requirement which presents a means of testing all numbers and selecting one, or finally the dependence can persist, but remain unknown" (quoted from Youshkevitch 1976, 77).

What is of note, and only apparently superfluous, is the list of the different modalities through which the function could be given which appears in this description. It is exactly this variety and diversity that makes up the basis for the formation of the abstract-theoretical function concept in a process of definition through abstraction.

The relativation of the role of any specific means used to define a function is expressed just as clearly in the quote from Dirichlet (1805–1859) by Felix Klein: "If, in an interval, every value  $x$  is assigned by *any means* (our italics) a definite value  $y$ , then  $y$  is a function of  $x$ " (Klein 1928, vol III).

That the concept should be determined by a mere input-output relationship, that is that it actually be identified by the fact that identical inputs or arguments result in identical outputs or function values, led to difficulties as well as to an uneasiness in mathematics that persisted throughout the 19th century. H. Hankel (1839–1873), for example, wrote in 1870, after he had reviewed the definition of a totally general function, "this purely nominal definition, which I will refer to as Dirichlet's from now on, ... is not sufficient for the needs of analysis, as functions of this type do not possess general properties, and therefore all relationships between function values and the different arguments no longer hold".

This abstract, conceptual view of the function just quoted transforms the function into a fully unknown object, as functions that are identical for a certain input can be totally different for a different one. It is not possible, as it were, to anticipate the "future" behaviour of such a function, that is the result of its application to arguments that have not yet been used.

For these reasons, which are also indicated by Hankel, the view of the function concept was inseparably connected with the "continuity principle". It is, so to speak, a pure operativity, undefinable without objective reference.

### References

- Agazzi, E.: 1974. "The rise of the foundational research in mathematics", *Synthese* 27, 7–26.
- Amsterdamski, S.: 1975. *Between Experience and Metaphysics*, Dordrecht, Reidel.
- Beth, E.: 1959. *The Foundations of Mathematics*, Amsterdam, North Holland.
- Beth, E. & Piaget, J.: 1966. *Mathematical Epistemology and Psychology*, Dordrecht, Reidel.
- Bochner, S.: 1969. *Eclosion and Synthesis*, New York, Springer.
- Bochner, S.: 1974. "Mathematical Reflections", *American Mathematical Monthly*, 827–852.
- Bolzano, B.: 1837. *Wissenschaftslehre. Versuch einer ausführlichen und größtenteils neuen Darstellung der Logik mit steter Rücksicht auf deren bisherige Bearbeiter*. In: Bolzano, B., *Gesamtausgabe*, edited by E. Winter, J. Berg, F. Kambartel, J. Louzil, B. van Rootselaar, vols. (I)11/1–(I)12/2, Stuttgart-Bad Cannstatt, Frommann Holzboog, 1985–1988.
- Bolzano, B.: 1975. "Erste Begriffe der allgemeinen Größenlehre (etwa 1830–1848)". In: Bolzano, B., *Gesamtausgabe*, edited by E. Winter, J. Berg, F. Kambartel, J. Louzil, B. van Rootselaar, vol. (II) A 7, 217–285, Stuttgart-Bad Cannstatt, Frommann Holzboog, 1975.
- Borzeskowski, H.-H. / Wahsner, R.: 1979. "Erkenntnistheoretischer Apriorismus und Einsteins Theorie", *Deutsche Zeitschrift für Philosophie* 27, 213–222.
- Bos, H. J. M.: 1984. "Arguments on motivation in the rise and decline of a

- mathematical theory; the 'Construction of Equations"', *Archive for History of Exact Sciences* 30, 331–380.
- Boutroux, P.: 1920. *L'idéal scientifique des mathématiciens*, Paris, Alcan.
- Casari, E.: 1974. "Axiomatical and Set-Theoretical Thinking", *Synthese* 27, 49–62.
- Cassirer, E.: 1910. *Substanzbegriff und Funktionsbegriff*, Berlin, Bruno Cassirer.
- Castonguay, Ch.: 1972. *Meaning and Existence in Mathematics*, Vienna, Springer.
- Engfer, H. J.: 1982. *Philosophie als Analysis: Studien zur Entwicklung philosophischer Analysiskonzeptionen unter dem Einfluß mathematischer Methodenmodelle im 17. und frühen 18. Jh.*, Stuttgart-Bad Cannstatt, Frommann Holzboog.
- Gaidenko, P.: 1981. "Ontologic foundation of scientific knowledge in 17th and 18th century rationalism". In: Jahnke, H. N. and Otte, M. (eds.), *Epistemological and social problems of the sciences in the early nineteenth century*. Dordrecht, Reidel.
- Gardies, L.: 1988/89. "La définition de l'identité d'Aristote à Zermelo", *Theoria* IV(10), 55–79.
- Gleick, J.: 1987. *Chaos. Making a New Science*, New York, Viking.
- Grassmann, H.: 1844. *Die Lineale Ausdehnungslehre*. Leipzig: All page references are taken from H. Grassmann. In: Engel, F. (ed.) *Gesammelte mathematische und physikalische Werke*, I.1, 1–319, Leipzig, Teubner, 1894.
- Grattan-Guinness, I.: 1970. *The Development of the Foundations of Mathematical Analysis from Euler to Riemann*, Cambridge, Mass., MIT Press.
- Hahn, H.: 1988. *Empirismus, Logik, Mathematik*, Frankfurt, Suhrkamp.
- Hankel, H.: 1870. "Untersuchungen über die unendlich oft oszillierenden und unstetigen Funktionen", *Math. Ann.* 20, 63–112.
- Hegel, G. W. F.: 1832. *Wissenschaft der Logik*. In: *Werke*, vol. 5, Frankfurt, Suhrkamp.
- Kant, I.: 1781. *Kritik der reinen Vernunft*, Riga: Hartknoch, J.F.: All page references are taken from I. Kant. *Werke*, (ed. by W. Weischedel), vol. 4, Frankfurt, Suhrkamp, 1956.
- Kitcher, Ph.: 1984. *The Nature of Mathematical Knowledge*, Oxford, Oxford UP.
- Klein, F.: 1928. *Elementarmathematik vom höheren Standpunkt*, 3 vols., Berlin, Springer.
- Krajewski, W.: 1977. *Correspondence Principle and the Growth of Science*, Dordrecht, Reidel.
- Lawvere, F. W.: 1992. "Categories of Space and of Quantity", this volume 14–30.
- Leibniz, G. W.: 1966. "Über das Kontinuitätsprinzip". In: *Hauptschriften zur Grundlegung der Philosophie* (edited by E. Cassirer), vol 1, 84–93. Hamburg, Meiner.
- Lovejoy, A. O.: 1936. *The Great Chain of Being*, Cambridge, Mass., Harvard UP.
- Mandelbrot, B.: 1983. *The Fractal Geometry of Nature*, New York.
- Markus, G.: 1986. *Language and Production*, Dordrecht, Reidel.

- Merz, J. T.: 1903. *History of European Thought in the 19th Century*, Edinburgh and London.
- Niiniluoto, I.: 1992, "Reality, Truth, and Confirmation in Mathematics – Reflections on the Quasi-Empiricist Programme", this volume 60-78.
- Otte, M. / Steinbring, H.: 1977. "Probleme der Begriffsentwicklung - Zum Stetigkeitsbegriff", *Didaktik der Mathematik* 1, 16-25.
- Otte, M.: 1984. "Komplementarität", *Dialektik* 8, 60-75.
- Otte, M.: 1989. "The Ideas of Hermann Grassmann in the Context of the Mathematical and Philosophical Tradition since Leibniz", *Historia Mathematica* 16, 1-35.
- Otte, M.: 1989a. "Die Auseinandersetzungen zwischen Mathematik und Technik als Problem der historischen Rolle und des Typus von Wissenschaft". In: S. Hensel et al., *Mathematik und Technik im 19. Jahrhundert in Deutschland*, Göttingen, 149-215.
- Otte, M.: 1990. "Arithmetic and Geometry: Some Remarks on the Concept of Complementarity", *Studies in Philosophy and Education* 10, 37-62.
- Otte, M.: 1990a. "Gleichheit und Gegenständlichkeit in der Begründung der Mathematik im 19. Jahrhundert - dargestellt am Beispiel der Auffassung von H. Grassmann, B. Bolzano und G. Frege", König, G. (ed.) *Konzepte des mathematisch Unendlichen im 19. Jahrhundert*, Göttingen. Vandenhoeck & Ruprecht.
- Peirce, Ch.S.: 1965. *Collected Papers*, 6 vols, Cambridge, Mass., Belknap Press.
- Peirce, Ch.S.: 1986. *Writings of Charles S. Peirce, A Chronological Edition*, vol. 4, Bloomington, Ind., Indiana UP.
- Peirce, Ch.S.: 1988. *Naturordnung und Zeichenprozeß*, Aachen, Alano Verlag.
- Rorty, R.: 1980. *Philosophy and the Mirror of Nature* (2nd ed.), Princeton, N.J., Princeton UP.
- Rota, G.-C. et al.: 1985. "On the Exterior Calculus of Invariant Theory", *Journal of Algebra* 96, 120-160.
- Rota, G.-C. et al.: 1986. *Discrete Thoughts*, New York-Berlin, Springer.
- Rota, G.-C. / Sherp, D.H. & Sokolowski, R.: 1988. "Syntax, Semantics and the Problem of the Identity of Mathematical Objects", *Philosophy of Science* 55, 376-386.
- Schelling, F. W. J.: 1802. *Vorlesungen über die Methode des akademischen Studiums*. In: Anrich, E. (ed.) *Die Idee der deutschen Universität*, 1-124, Darmstadt, Wissenschaftliche Buchgesellschaft, 1964.
- Schlick, M.: 1925/1971. *Allgemeine Erkenntnislehre*, Frankfurt.
- Simon, H. A.: 1969. *The Sciences of the Artificial*, Cambridge, Mass., MIT-Press.
- Strawson, P. F.: 1966. *The Bounds of Sense*, London, Methuen.
- Tuomela, R.: 1983. *Theoretical Concepts*, Vienna, Springer.
- Weyl, H.: 1968. *Gesammelte Abhandlungen*, 3 vols, Heidelberg, Springer.
- Youshkevitch, A. P.: 1976. "The Concept of Function up to the Middle of the 19th Century", *Arch. Hist. Exact Sci.* 16, 37-85.

# Turing's "Oracle": From Absolute to Relative Computability – and Back\*

SOLOMON FEFERMAN<sup>1</sup> (Stanford)

## 1. Introduction

What is offered here are some historical notes on the conceptual routes taken in the development of recursion theory over the last sixty years, and their possible significance for computational practice. These illustrate, incidentally, the vagaries to which mathematical ideas may be susceptible on the one hand, and – once keyed into a research program – their endless exploitation on the other.

At the hands primarily of mathematical logicians, the subject of *effective computability*, or *recursion theory* as it has come to be called (for historical reasons to be explained in the next section), has developed along several inter-related but conceptually distinctive lines. While this began with what were offered as analyses of the *absolute* limits of effective computability, the immediate primary aim was to establish *negative* results of effective unsolvability of various problems in logic and mathematics. From this the subject turned to refined classifications of unsolvability for which a myriad of techniques were developed. The germinal step, conceptually, was provided by Turing's notion of computability relative to an "oracle". At the hands of Post, this provided the beginning of the subject of *degrees of unsolvability*, which became a massive research program of great technical difficulty and combinatorial complexity. Less directly provided by Turing's notion, but implicit in it, were notions of *uniform relative computability*, which led to various important

---

\* Dedicated to S. C. Kleene for his many fundamental contributions to recursion theory.

<sup>1</sup> The ideas advanced in § 6 below were first presented in a special lecture entitled "Turing's 'oracle'" at Carnegie-Mellon University, July 1, 1987. I wish to thank T. Fernando, M. Lerman, P. Odifreddi and W. Sieg for their useful comments on a draft of this paper.

*theories of recursive functionals*. Finally, the idea of computability has been relativized by extension, in various ways, to more or less *arbitrary structures*, leading to what has come to be called *generalized recursion theory*. Marching in under the banner of degree theory, these strands were to some extent woven together by the recursion theorists, but the trend has been to pull the subject of effective computability even farther away from questions of actual computation. The rise in recent years of *computation theory* as a subject with *that* as its primary concern, forces a reconsideration of notions of computability both in theory and in practice. Following the historical sections, I shall make the case for the primary significance for practice of the various notions of relative (rather than absolute) computability, but *not* of most methods or results obtained thereto in recursion theory.

While a great deal of attention has been paid in the literature to the early history of recursion theory in the '30s, and to the grounds for the so-called Church-Turing Thesis as to absolute effective computability, hardly any has been devoted to notions of relative computability. The historical sketch here is neither definitive nor comprehensive; rather, the intention is to mark out the principal conceptual routes of development with the end purpose of assessing their significance for computational practice. Nor is any claim made as to the "right" generalization, if any, of computability to arbitrary structures. However, the time is ripe for a detailed historical study of relative computability in all its guises and for the assessment of proposed generalizations of the Church-Turing Thesis.

## 2. "Absolute" Effective Computability

### 2.1. Machines and Recursive Functions

The history of the development of what appeared to be the most general concept of effective computability and of the Church-Turing Thesis thereto, is now generally well known. Cf. Kleene 1981, Davis 1982, Gandy 1988 and (especially for Turing's role) Hodges 1983.

By 1936, several very different looking proposals had been made for the explication of this notion:  $\lambda$ -definability (Church), general recursiveness (Herbrand-Gödel), and computability by machine (Turing and, independently, Post). These were proved in short order to lead to co-extensive classes of functions (of natural numbers). In later years, still further notions leading to the same class of functions were introduced; of these we mention (for purposes below) only computability on register machines introduced by Shepherdson and Sturgis in 1963.



The definition generally regarded as being the most persuasive for Church's Thesis is that of computability by Turing machines. This is not described here because of its familiarity and because of its remove from computational practice. The register machine model of Shepherdson-Sturgis is closer to the actual design of computers and even leads to a "baby" imperative-style programming language. As with all the definitions mentioned above, it makes idealized assumptions to the effect that work-space and computation-time are unlimited. Each register machine is provided with a finite number of memory locations  $R_i$ , called registers, each of which has an unlimited capacity in the sense that arbitrarily long sequences of symbols and numbers can be stored in them. Here we restrict attention to computation over the set  $N = \{0, 1, 2, 3, \dots\}$  of natural numbers, to which more general symbolic computation can be reduced. Certain registers  $R_1, \dots, R_n$  are reserved for inputs (when computing a function  $f: N^n \rightarrow N$ ), and one (say  $R_0$ ) is reserved for the output; the other registers provide memory locations for computations in progress. A program is given by a sequence of instructions  $I_0, \dots, I_m$  of which  $I_0$  is the initial instruction and  $I_m$  is the final or HALT instruction. The active instructions  $I_j$  ( $0 \leq j < m$ ) are of one of the following forms: (i) increment the contents  $r_i$  of  $R_i$  by 1, (ii) decrement by 1 (if different from 0), (iii) set the contents of  $R_i$  to 0, and (iv) test to see if  $r_i$  is 0 and then branch to one or another instruction depending on the answer. These are symbolized respectively by

- (i)  $r_i := r_i + 1$ ,
- (ii)  $r_i := r_i - 1$ ,
- (iii)  $r_i := 0$ ,
- (iv) if  $r_i = 0$  go to  $I_k$ , else to  $I_l$ .

A function  $f: N^n \rightarrow N$  is computable by such a machine if when we load any natural numbers  $x_1, \dots, x_n$  as input in  $R_1, \dots, R_n$ , resp., and begin with instruction  $I_0$ , the output  $f(x_1, \dots, x_n)$  will eventually appear in  $R_0$  at the HALT state  $I_m$ . As mentioned above, it has been shown that the class of register computable functions is the same as that of Turing computable functions.

Returning to the situation in the latter part of the 30's, the results establishing the (extensional) equivalence of the various proposed definitions of effective computability bolstered Church's Thesis. Church himself had announced this in terms of the Herbrand-Gödel notion of *general recursiveness*. The fact that many effectively computable functions in practice were given by recursive defining equations led logicians to treat effective computability and recursiveness as interchangeable notions. Thus the subject has come to be called *recursive function theory*, or simply *recursion theory*.

As a result especially of Turing's analysis, Church's Thesis has gained

almost universal acceptance; the case for it was assembled in Kleene 1952 (especially in §§ 62, 63 and 70). For a more recent analysis, see Gandy 1980,<sup>2</sup> and for a comprehensive survey of the literature on this subject, see Odifreddi 1989, § I. 8. Many would agree with Gödel that the importance of general recursiveness (or Turing computability) is that " ... with this concept one has for the first time succeeded in giving an *absolute* definition of an interesting epistemological notion, i.e., one not depending on the formalism chosen". (Italics mine; quotation from the 1946 article appearing in Gödel 1989).

## 2.2. Partial Recursive Functions

In his paper 1938, Kleene made an essential conceptual step forward by introducing the notion of *partial recursive function*. This can be explained in terms of any of the models of effective computability mentioned above, in particular by means of the Turing or register machine approaches. Each instruction sequence or program  $I = I_0, \dots, I_m$  may be coded by a natural number  $i$ , and then  $M_i$  is used to denote the corresponding machine. For simplicity, consider functions of one argument only; given an arbitrary such  $x$ ,  $M_i$  need not terminate at that input. If it does, we write  $M_i(x)$  for the output value and say that  $M_i(x)$  is defined. This then determines a function  $f$  with domain  $\text{dom}(f) = \{x \mid M_i(x) \text{ is defined}\}$ , whose value for each  $x$  in  $\text{dom}(f)$  is given by  $M_i(x)$ . A function is said to be partial recursive just in case it is one of these  $f$ 's.

Kleene established a *Normal Form Theorem* for partial recursive functions as follows (adapted to the machine model). Let  $C(i, x, y)$  mean that  $y$  codes a terminating computation on  $M_i$  at input  $x$ ; it may be seen that  $C$  is an effectively decidable relation (in fact, in the subclass of primitive recursive relations). Moreover, the function  $U(y)$  which extracts the output of  $y$  when  $y$  represents a terminating computation, and is otherwise 0, is also (primitive) recursive. Hence, for partial recursive  $f$  determined by  $M_i$  as above, we have

- (1) (i)  $\text{dom}(f) = \{x \mid (\exists y) C(i, x, y)\}$ , and
- (ii)  $f(x) = U((\text{least } y) C(i, x, y))$  for each  $x$  in  $\text{dom}(f)$ .

Moreover, every function defined in this way is partial recursive. One may further observe from this result that if we define  $g(i, x)$  for all  $i, x$  by

- (2)  $g(i, x) = U((\text{least } y) C(i, x, y))$  whenever  $M_i(x)$  is defined,

then  $g$  is a partial recursive function of two arguments which ranges for  $i = 0$ ,

---

<sup>2</sup> And the still more recent critical discussion provided in the dissertation Tamburrini 1987.

1, 2, ... (as a function of  $x$ ) over all partial recursive functions of one argument. This is what is called the *Enumeration Theorem* for partial recursive functions. Now since  $g$  itself is computable by a machine  $M$ , we have the consequence – already recognized by Turing in 1937 – that there is a *universal computing machine*  $M$ , which can simulate the behavior of any other  $M_i$  by acting at input  $(i, x)$  as  $M_i$  acts at input  $x$ . This in turn may be considered to provide the conceptual basis for *general purpose computers*, which can store "software" programs  $I$  in the form of data " $i$ " for any particular application.

### 2.3. Effectively Unsolvable Problems and the Method of Reduction

A number of questions had been raised in the period 1900–1930 concerning the uniform effective solvability of certain classes of mathematical problems. Outstanding among these were:

- (i) *Diophantine equations*. To decide, if possible, whether a polynomial equation with integer coefficients has integer solutions ("Hilbert's 10th problem").
- (ii) *Word problem for groups*. To decide, if possible, whether two words in a finitely presented group represent the same element or not.
- (iii) *Entscheidungsproblem*. To decide, if possible, whether a given formula in the first-order predicate calculus (1st order PC) of logic is satisfiable or not.

While partial progress was made on each of these problems for specific cases, the general problems resisted positive solution. In particular, for (iii), initial optimism was tempered by the famous incompleteness results of Gödel in 1931 (cf. the collection 1986) in which it was shown, among other things, that for any sufficiently strong and correct formal system  $S$  one can produce a formula  $A_S$  of the 1st order PC such that  $A_S$  is satisfiable, but that fact is not provable in  $S$ . Hence if there were a decision method  $D$  for satisfiability in the 1st order PC, no such  $S$  could verify that  $D$  works.

But in order to obtain definitive negative results concerning these and similar problems, one would need a precise and completely general definition of effective method. This would be analogous to supplying a general definition of ruler-and-compass construction in order to show the non-constructibility of the classical geometric problems (angle trisection, duplication of the cube, etc.), or of solvability by radicals in order to demonstrate the nonsolvability in such terms of various algebraic equations (of 5th degree and beyond). In the case of effective computability and effective decidability, that is just what was supplied by the definitions described in § 2.1, according to the Church-Turing Thesis. And, indeed, Church and Turing used this to establish the effective un-

solvability of the *Entscheidungsproblem*. Their thesis is implicitly taken for granted in the following, where we analyze one aspect of their proofs.

Given a set  $A$  of natural numbers, the *membership problem* for  $A$  is the question whether one has an effective method to decide, given a natural number  $x$ , whether or not  $x$  belongs to  $A$ . The characteristic function  $c_A$  of  $A$  is the function defined by  $c_A(x) = 1$  if  $x$  is in  $A$ , and 0 otherwise. The membership problem for  $A$  is effectively solvable just in case  $c_A$  is effectively computable. If this holds we say that  $A$  itself is computable or recursive, while if  $A$  is not recursive, its membership problem is said to be effectively unsolvable. The very first example of such a problem provided by Turing was that of the *Halting Problem* (H. P.) for Turing Machines, which is readily adapted to register machines. This is to decide, given  $i$  and  $x$ , whether or not  $M_i(x)$  is defined. The *Diagonal H. P.* is the question whether  $M_x$  terminates at input  $x$ , and it is represented by the set  $K = \{x \mid M_x(x) \text{ is defined}\}$ . Now one easily shows by a diagonal argument that  $K$  cannot be recursive. For if it were, its characteristic function  $c_K$  would be recursive. But then so would be the function  $d(x) = (M_x(x) + 1)$  if  $x$  is in  $K$ , and 0 otherwise. Now since  $d$  is recursive, it is computed by some specific machine  $M_i$ , i.e.  $d(x) = M_i(x)$  for all  $x$ . Then, in particular,  $d(i) = M_i(i)$ , contradicting  $d(i) = M_i(i) + 1$ .

The general halting problem is represented by the set  $H = \{(x, y) \mid M_x(y) \text{ is defined}\}$ . Clearly,  $x$  is in  $K$  just in case  $(x, x)$  is in  $H$ . If  $H$  were recursive, then  $K$  would be recursive, contrary to what has just been shown. The general situation here is given by the notion of *many-one reduction*,  $A \leq_m B$ . This is defined to hold just in case there is a recursive function  $f$  such that for all  $x$

(1)  $x$  is in  $A$  if and only if  $f(x)$  is in  $B$ .

(It is called "many-one" because the function  $f$  might give the same value for many arguments). We have the trivial result:

*Theorem.* If  $A \leq_m B$  and  $B$  is recursive, then  $A$  is recursive. Hence if  $A$  is not recursive,  $B$  is not recursive.

In essence Turing established the negative result for the *Entscheidungsproblem* (in his paper 1936–37) by taking  $S = \{x \mid x \text{ is the number of a formula in the 1st order PC which is satisfiable in some model}\}$  and showing  $K \leq_m S$  where  $K$  is the diagonal halting problem. (The argument in Church 1936 makes use instead of reduction from an effectively unsolvable problem in the  $\lambda$ -calculus).

The relation  $\leq_m$  of many-one reduction is one of the most widely applied in practice for effective unsolvability results. Eventually, Hilbert's 10th problem

and the word problem for groups were shown to be effectively unsolvable<sup>3</sup> by reducing the problem  $K$  to them (through a long chain of arguments). However, it is not in principle the most general relation  $\leq$  between the sets  $A$  and  $B$  which will allow one to conclude:

(2) If  $A \leq B$  and  $B$  is recursive, then  $A$  is recursive.

For example, we could take  $\leq$  to mean that there is a *truth-table reduction* of  $A$  to  $B$ , i.e. that membership of an  $x$  in  $A$  is determined by a propositional combination of statements of the form  $f_i(x)$  in  $B$ , for  $f_i$  recursive; in such a case we write  $A \leq_{tt} B$ . The most general concept of effective reduction of the membership problem for one set to another is wider still. This was provided by Turing's notion of computation relative to an "oracle", to which we now turn.

### 3. Relative Effective Computability over the Natural Numbers

#### 3.1. Turing's "Oracle" and Turing Reducibility

The germinal idea of computability relative to an "oracle" was introduced almost as an aside in *Turing 1939*, a paper based on Turing's Ph. D. thesis at Princeton University under the direction of Church. The story of how Turing came to do graduate work in Princeton and of the outcome of his studies is told in *Hodges 1983*, pp. 90–146, and again in *Feferman 1988*, where the contents of the thesis publication are analyzed in some detail. Turing's dissertation work concerned the concept of ordinal logics, which were introduced in an attempt to overcome Gödel's incompleteness results by the iterated (finite and transfinite) adjunction to each formal system (or logic)  $S$  accepted in the process, of such statements as  $Con_S$  expressing the consistency of  $S$ , shown by Gödel to be unprovable in  $S$  though informally recognized to be true. Turing's aim was to obtain completeness for two-quantifier ( $Q_2$ ) statements of arithmetic, i.e., those of the form  $(x)(\exists y) R(x,y)$  with  $R$  recursive, and in this he was only partially successful. Now the section of the published dissertation (1939) in which Turing introduces the "oracle" notion is a brief one, § 4, whose main aim is to produce a mathematical problem which is not in  $Q_2$  form. The existence of non- $Q_2$  definable sets is of course trivial by a cardinality or diagonal argument, but presumably Turing wanted to

---

<sup>3</sup> Through the work, respectively of Davis, Putnam, J. Robinson and Matijasevich for the former problem, and (with successive improvements) Novikoff, Boone, Britton and Higman for the latter.

produce something with more concrete mathematical content. He begins by saying: "Let us suppose that we are supplied with some specified means of solving number-theoretic  $[Q_2]$  problems; a kind of oracle as it were ... With the help of the oracle we could form a new kind of machine (call them  $O$ -machines), having as one of its processes that of solving a given number-theoretic problem". Turing then shows (loc. cit.) more precisely how to define computability by an  $O$ -machine and, by a direct extension of his argument in 1936–37, that (in effect) the Halting Problem for  $O$ -machines is not decidable by an  $O$ -machine and hence is outside the class of  $Q_2$  problems assumed to be decided by  $O$ .

Turing did nothing further with this idea and it was not until *Post 1944* that it began to be taken as the basis for a systematic investigation. To begin with, the idea of an  $O$ -machine is directly generalized to that of a  $B$ -machine for any set  $B$ . In the register machine model, one simply adds to the basic instructions, ones of the form:

(1)  $r_j := 1$  if  $r_k$  is in  $B$ , else  $r_j := 0$ .

Given a list  $I$  of instructions for computation relative to a set in the sense just explained, and given a specific set  $B$ , if the computation by the machine  $M$  determined by  $I$  terminates at any given  $x$ , we write  $M^B(x)$  for the output. If for each  $x$ ,  $M^B(x) = 0$  or  $1$ , then  $M^B$  is the characteristic function of a set  $A$  with

(2)  $x$  in  $A$  if and only if  $M^B(x) = 1$ .

In this case we say  $A$  is *Turing computable from  $B$*  or *Turing reducible to  $B$*  and write  $A \leq_T B$ .

It is not hard to see that

*Lemma* (i)  $A \leq_m B \Rightarrow A \leq_u B$  ; (ii)  $A \leq_u B \Rightarrow A \leq_T B$ .

Moreover, the arguments for the Church-Turing Thesis lead one strongly to accept a relativized version: (C-T) <sup>(\*)</sup>  $A$  is effectively computable from  $B$  if (and only if)  $A \leq_T B$ .

Thus Turing reducibility gives the most general concept of relative effective computability.

The relation of computability of one function from another is even more simply defined by an extension of the definition of register computability. Given a function  $g: N \rightarrow N$  we add to the four previous register instructions of § 2.1. (1), instructions of the form

(3)  $r_j := g(r_k)$ ,

whose meaning is to set the content of register  $R_j$  to be  $g(r_k)$  where  $r_k$  is in

$R_k$ . Then  $f \leq_T g$  if and only if  $f$  is register computable from  $g$  in this expanded sense. Note that for sets  $A$  and  $B$  we have

(4)  $A \leq_T B$  if and only if  $c_A \leq_T c_B$ .

where  $c_A, c_B$  are the characteristic functions of  $A$  and  $B$ , resp. Thus there is really only one basic notion of relative effective computability involved, namely that for functions relative to functions. However, *Post 1944* concentrated on the relation  $A \leq_T B$ , since the classical effective (un)solvability problems concerned the membership problem for sets, and in particular for a special class of sets, the recursively enumerable ones, to which we turn next.

### 3.2. Recursively Enumerable Sets, Degrees of Unsolvability, and Post's Problem

A set  $A$  of natural numbers is said to be *recursively enumerable* (r.e.) if it is the range of a recursive function, i.e.,

(1)  $A = \{ f(0), f(1), \dots, f(x), \dots \}$

for some recursive  $f$ , or if  $A$  is empty. The empty set is included as a limiting case, so that each r.e. set may also be written in the form

(2)  $x$  in  $A$  if and only if  $(\exists y) R(x, y)$

with  $R$  a recursive relation, and conversely. Clearly every set  $A$  of the form (1), as well as the empty set, is of the form (2). For the converse, one uses a recursive pairing function  $p$  with inverses  $p_0, p_1$  so that  $p(p_0(x), p_1(x)) = x$  for each  $x$ . Then if  $A$  is non-empty, say  $x_0$  in  $A$ , and (2) holds, let  $f(x) = p_0(x)$  if  $R(p_0(x), p_1(x))$  and  $f(x) = x_0$  otherwise, so that (1) holds. Note that there may be repetitions in (1), so a non-empty r.e. set could be finite.

Every recursive set  $A$  is recursively enumerable, as we see from (2) by taking  $R(x, y)$  to be:  $x$  in  $A$  &  $y = 0$ . However, the converse is not true: the diagonal halting set  $K$  is recursively enumerable but not recursive. The latter was argued above; to see that  $K$  is r.e. one simply uses the fact that  $K = \{ x \mid (\exists y) C(x, y) \}$ , in the symbolism of § 2.2 above.

It is not hard to see that all the classical problems of effective (un)solvability mentioned in § 2.2 concern r.e. sets. For example, Diophantine sets are those in the form

(3)  $D_{p,q} = \{ x \mid (\exists y_1) \dots (\exists y_n) (p(x, y_1, \dots, y_n) = q(x, y_1, \dots, y_n)) \}$

where  $p, q$  are polynomials with coefficients in  $N$ , and these reduce to the

form (2) by combining the prefix  $(\exists y_1) \dots (\exists y_n)$  into a single  $(\exists y)$  using the pairing functions. The *Entscheidungsproblem* is a special case of the general *decision problem* for formal systems  $S$ , to decide whether or not a given formula  $A$  is provable from  $S$ . Using Gödel-numbering of formulas and derivations, this reduces to the question whether the set  $Prov_S$  given by

$$(4) \text{ } Prov_S = \{ x \mid (\exists y) \text{ } Proof_S(x, y) \}$$

is recursive. But  $Prov_S$  is r. e. whenever  $S$  is recursive, since then  $Proof_S$  is recursive. The *word problem* for groups is the question whether an equation between words in a finitely presented group is provable from the defining equations of the presentation and the group axioms by the rules of equality, and this leads again to sets of the form (2); similarly for other algebraic systems. Not all natural effectiveness problems are recursively enumerable. The first (*prima facie*) more complicated problem is whether any given unary partial recursive function determined by  $M_i$  is total, i. e., whether  $i$  belongs to the set

$$(5) \text{ } Tot = \{ z \mid (x) (\exists y) C(z, x, y) \}.$$

One easily sees that every r. e. set  $A$  is reducible to the Halting Problem, as follows. For  $A$  r. e. in the form (2), consider the function

$$(6) \text{ } f(x) = (\text{least } y) R(x, y).$$

Then  $f(x)$  is defined just in case  $(\exists y) R(x, y)$ , so  $\text{dom}(f) = A$ . Moreover,  $f$  is partial recursive, so  $f(x) = M_i(x)$  for some  $i$  and all  $x$  in  $\text{dom}(f)$ . Hence we have  $x$  in  $A$  if and only if  $M_i(x)$  defined, i. e. if and only if  $(i, x)$  is in  $H$ . In other words

$$(7) \text{ } A \leq_m H.$$

A little more detailed argument is required to prove the following:

*Lemma.* For each r. e. set  $A$ , we have  $A \leq_m K$ .

(For a proof, cf. Kleene 1952, p. 343). Thus  $K$  is what Post called *complete* for the class of r. e. sets, i.e., it is r.e. and every r.e. set is reducible to it.

Now Post defined a set  $A$  to have *equal or lower degree of unsolvability* than  $B$  if  $A \leq_T B$  and  $A$  to have the *same degree of unsolvability* as  $B$  if both  $A \leq_T B$  and  $B \leq_T A$ . The latter is an equivalence relation between sets of natural numbers; the equivalence class of a set  $A$  is called its *degree of unsolvability* and denoted  $\text{deg}(A)$ . We use letters  $a, b, \dots$  to range over degrees of unsolvability. Given  $a = \text{deg}(A)$ ,  $b = \text{deg}(B)$  we take  $a \leq b$  just in case  $A \leq_T B$ , and  $a < b$  if  $a \leq b$  but  $a \neq b$ , i.e., if  $A \leq_T B$  but not  $B \leq_T A$ . Note



that  $a = b$  iff  $a \leq b$  &  $b \leq a$ .

Let  $0 = \deg(N)$ . Since  $N$  is recursive, we have  $N \leq_T A$  for any set  $A$ , hence

(8)  $0 \leq a$  for all degrees  $a$ .

It seems anomalous to call  $0$  a degree of unsolvability since for any  $A$  with  $0 = \deg(A)$  we have  $A \leq_T N$ , hence  $A$  is recursive, i.e., is effectively decidable. However,  $0$  is a limiting case of the degrees:

(9) if  $0 < \deg(A)$ , then  $A$  is not recursive.

Let  $0' = \deg(K)$ . Then by the Lemma above, we have

(10) for each r.e. set  $A$ ,  $\deg(A) \leq 0'$ .

As we observed in § 2.2, every r.e. set  $A$  met in practice and which has been shown to be non-recursive, has been shown to be so by a chain of reductions leading to  $K \leq_m A$ , and so, of course  $K \leq_T A$  in all such cases. Post raised the question in 1944 whether this must be so in general. That is, he asked:

*Post's problem.* Do there exist r.e. sets  $A$  with  $0 < \deg(A) < 0'$ ?

As Post put it (1944, pp. 289–290): "Our whole development largely centers on the single question whether there is, among these problems [for recursively enumerable non-recursive sets] a lower degree of unsolvability than that [of the highest degree,  $\deg(K)$ ], or whether they are all of the same degree of unsolvability". After crediting *Turing 1939* for the basic idea of reducibility and for establishing in effect that for any set  $A$  there is one of a higher degree of unsolvability, Post goes on to say: "While [Turing's] theorem does not help us in our search for that lower degree of unsolvability, his formulation makes our problem precise. It remains a problem at the end of this paper. But on the way we do obtain a number of special results, and towards the end obtain some idea of the difficulties of the general problem".

The "special results" that Post refers to here concerned the existence of lower degrees of unsolvability with respect to the reducibility relations  $\leq_m$  and  $\leq_u$ . Thus he produced the existence of a non-recursive set  $S$  that he called "simple", such that  $K$  is not many-one reducible to  $S$ ; however,  $K \leq_u S$  (if one allows unbounded truth-tables). Then Post produced a non-recursive r.e. set  $S^*$  that he called "hyper-simple", such that  $K$  is not truth-table reducible to  $S^*$ ; however,  $K \leq_T S^*$ , so Post asked whether there might not exist "hyper-hyper-simple" sets which evade this reduction. As the constructions became combinatorially more and more complicated, the difficulty of Post's problem became evident. Towards the end of his 1944 paper, Post said that: "As a result we are left completely on the fence as to whether there exists a recursively enumerable [non-recursive] set of positive integers of absolutely lower degree

of unsolvability than the complete set  $K$ , or whether, indeed, all recursively enumerable sets of positive integers with recursively unsolvable decision problems are absolutely of the same degree of unsolvability".

### 3.3. *The Solution of Post's Problem and the Flowering of Degree Theory*

The first results pushing toward a solution of Post's problem were obtained by Kleene and Post in their joint publication 1954. This also led to a basic bifurcation in the subject of degrees of unsolvability, already implicit in Post's remarks quoted above. Namely, one can consider the relation  $A \leq_T B$  without restriction on the way  $A, B$  may be defined. Post's problem had concerned the relation  $\leq_T$  restricted to r.e. sets, but one may consider it for *any* sets  $A, B$ , one or both of which might not be r.e. As pointed out by Post, Turing's original construction (in his 1939 paper) in effect associated with each set  $A$  another set  $A'$  such that

$$(1) \deg(A) < \deg(A'),$$

by taking  $A'$  to be the diagonal halting problem  $K^A$  relativized to  $A$ , i.e.,

$$(2) A' = K^A = \{x \mid M_x^A(x) \text{ is defined}\}.$$

This  $K^A$  is obtained from a predicate (primitive) recursive in  $A$  by prefixing one numerical (existential) quantifier. In degree notation,

$$(3) a < a', \text{ where } a' = \deg(K^A) \text{ for } a = \deg(A).$$

In particular,

$$(4) 0 < 0' < 0'' < \dots$$

But this is only a crude classification of the degrees of unsolvability of arithmetically definable sets. What Kleene and Post showed in 1954 (among other things) is that between each of the inequalities  $a < a'$  in (3) there are infinitely many other degrees, in fact, there is a subset  $D$  of  $\{d \mid a < d < a'\}$  such that  $D$  is densely ordered by the  $<$  relation. Naturally, if any  $d$  with  $0 < d < 0'$  were the degree of an r.e. set this would be the solution of Post's problem. However, the Kleene-Post proof was not sufficiently effective to establish this, and their set  $D$  consists entirely of non-r.e. degrees.

While considerable effort was devoted by a number of logicians to Post's problem in the dozen years following its publication, there was no breakthrough until 1956, when the problem was solved, independently, by R. Friedberg and A. A. Muchnik (see the references in *Rogers 1967*). At the time, Friedberg was a 20-year-old senior at Harvard University, and Muchnik

was hardly any older. Friedberg had learned of the problem while taking a course on recursion theory taught by Hartley Rogers of M. I. T. Friedberg's and Muchnik's solution established the existence of *two r.e. sets,  $A, B$  of incomparable degree*, i.e. for which

(5) (i)  $A, B$  are r.e. , and (ii) neither  $\deg(A) \leq \deg(B)$  nor  $\deg(B) \leq \deg(A)$ .

Thus, for  $a = \deg(A)$  and  $b = \deg(B)$  we have

(6)  $0 < a < 0'$  and  $0 < b < 0'$ ;

for, if either one of  $a, b$  were equal to  $0$  or  $0'$ , (5) (ii) would be false.

A special new technique was introduced by Friedberg (and Muchnik) for the result (5), called the *priority method*. The sets  $A, B$  are constructed in stages and at each stage only a finite number of membership and non-membership relations have been set down though only tentatively. Each stage is devoted to finding for specific  $i$  an  $n$  such that  $c_A(n) \neq M_i^B(n)$  (and similarly for  $A, B$  reversed). If this is successful, one puts  $n$  in  $A$  if  $c_A(n) = 1$ , otherwise out of  $A$ ; and similarly for  $B$ . However, when we thus enlarge (the characteristic function of)  $A$ , it may turn out that this affects the value assigned to  $M_j^A(m)$  for some  $j, m$  at a previous stage. By assigning priorities to the actions, it is shown that at most a finite number of changes can take place for each  $i$ ; this kind of argument is thus often called the *finite-injury method*. The argument for (5) is not long (it only took three pages to communicate it adequately in *Friedberg 1957*)<sup>4</sup> but its novelty, ingenuity and the circumstances of its discovery were stunning, both to logicians and to a mass audience. (For example, the news was reported in *Time* magazine for March 19, 1956, p. 83). In the next few years, Friedberg made several other interesting applications of the priority method, but after that left the field completely. The work of Friedberg and Muchnik opened the flood gates to the development of the *theory of degrees of unsolvability* (or simply *degree theory*, as it is often called) both for r.e. degrees and degrees of arbitrary sets. In a survey a decade later, Simpson (1977, p. 632) said that "...for many years now, degree theory has been one of

<sup>4</sup> See also the three-page proof in *Rogers 1967*, pp. 163–166. Regarding novelty, however, Rogers says (loc. cit., p. 163) that "in their initial presentations, both Friedberg and Muchnik built on earlier ideas and results of Kleene and Post". As it turned out later, the priority arguments are not essential for the solution of Post's problem, due to work of Kucera. See *Odifreddi 1989*, Ch. III for this and other treatments more in the spirit of *Post 1944*; that also gives interesting background on Post's problem, beginning with Post's own preliminary ideas on undecidability in the 1920's.

the most technical and highly developed parts of mathematical logic. There are literally hundreds of papers in the literature, all devoted exclusively to degrees. The standard of originality in these papers is very high. Although certain ideas recur, the variety of methods employed is enormous".

The effect this development had on recursion theory is evident from a comparison of the expositions of the subject in Kleene 1952, Part III with that of Rogers 1967 or Odifreddi 1989, which are dominated by notions and results concerning the reduction relations  $\leq_m$ ,  $\leq_u$ ,  $\leq_T$  and the associated degrees. The first monograph to be devoted entirely to the subject of degree theory and which contained many important new contributions was that of Sacks 1963. (In particular, that extended the priority method to permit *infinite injury* arguments). More recently, the books of Lerman 1983 and Soare 1987 serve to bring graduate students and young researchers to the forefront of research, emphasizing, respectively, degrees of arbitrary sets and degrees of r.e. sets; the bibliography of the latter book contains on the order of 600 entries. While the subject of recursion theory has also developed along other lines, some of which will be seen in the following, none matches degree theory for its level of difficulty and for the complexity of the arguments involved (as suggested by the quotation above from Simpson). It is for this reason that one sometimes hears of the periods of development of recursion theory divided into "pre-Post" and "post-Post" or "pre-Friedberg" and "post-Friedberg". On the face of it, the results of degree theory are irrelevant to computation theory, because they concern effectively unsolvable problems. However, they may be suggestive of analogous results for degrees of complexity of (effective) algorithms; we shall return to the possible connection in § 6 below.

#### 4. Uniform Relative Computability over the Natural Numbers

##### 4.1. Relative Computation Procedures and Partial Recursive Functionals

In Turing's conception of relative computability, the "oracle" is queried for information about some fixed set. That is, given  $B$ , we try to find out for various  $A$  whether or not  $A \leq_T B$ , i.e., whether or not we can find an  $i$  such that  $A = M_i^B$ . Put in terms of functions the question is, for given  $f, g$  whether or not there exists  $i$  such that  $f = M_i^g$ .

Now when dealing with the idea of *uniform relative computability* we shift attention from fixed  $f, g$  and possible computation instructions (coded by)  $i$  connecting them to fixed  $i$  and the effect of varying  $g$  in  $M_i^g$  on  $f = M_i^g$ . That is, we fix a *relative computation procedure* and consider its effect on the functions which result when we vary the functions to which it applies. In general,

such a procedure  $F$  will take us from several different functions  $g_1, \dots, g_m$  to a new function  $f$  in the form

$$(1) f = F(g_1, \dots, g_m).$$

Such  $F$  are called *functionals*. Concrete examples of *effective functionals* are provided by composition and minimalization, in the form:

$$\begin{aligned} (\text{Comp}) \quad & f(x) = g_1(g_2(x), g_3(x)) \\ (\text{Min}) \quad & f(x) = (\text{least } y) [g(x, y) = 0]. \end{aligned}$$

In (Comp),  $f = F_1(g_1, g_2, g_3)$ , and in (Min),  $f = F_2(g)$ . To begin with, these schemata were conceived of as applying to *total* functions  $g, g_1, g_2, g_3$  on  $N$  and to lead to *total* functions  $f$ , provided for (Min) that  $(x) (\exists y) g(x, y) = 0$ . In that schema, if we assume  $g$  is total but don't know whether there is always a  $y$  with  $g(x, y) = 0$ , we can only conclude that  $f$  is partially defined. Then the '=' sign in (Min) must be replaced by ' $\simeq$ ', which means that if either side is defined, so is the other, and they are equal. If we go further to allow partial functions to appear on the right side of these equations then we must also replace '=' by ' $\simeq$ ' in Comp and in similar examples. Thus partial recursive functions are allowed in general to operate on partial functionals. When written in the form (1), the '=' sign is appropriate, since one obtains a well-determined partial function  $f$  from  $(g_1, \dots, g_m)$  by application of  $F$ . However,  $f$  itself might be nowhere defined. In general then, we must write

$$\begin{aligned} (2) f(x_1, \dots, x_n) &\simeq (F(g_1, \dots, g_m))(x_1, \dots, x_n), \text{ for which we also write} \\ (3) f(x_1, \dots, x_n) &\simeq F(g_1, \dots, g_m; x_1, \dots, x_n). \end{aligned}$$

The notion of partial recursive functional, which has yet to be explained, is easily defined by means of any of the basic approaches to recursion theory. For simplicity, consider  $n = m = 1$  in the above. In the register machine approach, the partial recursive functionals  $F$  are all those definable in the form

$$(4) F(g) = M_i^g$$

for some fixed  $i$ , so that  $F(g; x) \simeq M_i^g(x)$  for all  $g, x$ . In other words, we are interpreting the instruction set given by  $i$ , including rules of the form,  $r_j: = g(r_k)$ , to apply to *variable* partial functions  $g$ .

#### 4.2. The Recursion Theorems

The notion of partial recursive functional was introduced by Kleene in 1950. He said that he arrived at this "by considering Turing's computation by a machine having access to an oracle, but with the rules governing the machine ...

fixed, *varying the oracle* so that she answers for one or another value of a ... function variable ..." (Kleene 1981, p. 64, italics mine). The basic theory of partial recursive functionals was given its first systematic exposition in Kleene 1952, Ch. XII. To begin with, the following properties are easily established, where again for simplicity we consider the case  $n = m = 1$  and write  $F(g;x)$  for  $F(g)(x)$ .

*Lemma.* Suppose  $F$  is a partial recursive functional.

- (i) (Monotonicity) If  $g_1$  is contained in  $g_2$ , then  $F(g_1)$  is contained in  $F(g_2)$ .
- (ii) (Continuity) If  $F(g;x) \simeq y$ , then for some finite  $h$  contained in  $g$  we have  $F(h;x) \simeq y$ .
- (iii) (Effectiveness) If  $g$  is partial recursive, then  $F(g)$  is partial recursive.

The fundamental result obtained by Kleene for these functionals is the following (op. cit., p. 348):

*The Recursion Theorem (Functional Form).* For any partial recursive functional  $F(g;x)$  there is a least solution  $f$  to the equation  $f(x) \simeq F(f;x)$ . Moreover, that  $f$  is partial recursive.

The proof simply takes  $f =$  the union of the  $g_n$ 's, where  $g_0$  is the empty function and  $g_{n+1} = F(g_n)$ . Then  $f$  is the least solution of the equation  $f = F(f)$  by the monotonicity and continuity properties of  $F$ . To prove that  $f$  is partial recursive, one makes use of the effectiveness property.

The Recursion Theorem, and another, index form (already established by Kleene in 1938) have many applications in recursion theory. The reason for the importance of these results is that they apparently give the most general effective versions of defining a function over the natural numbers recursively, i.e., in terms of itself. While the index form seems to be applied more often in practice, the functional form may be considered more fundamental. For, it is expressed in *intrinsic* terms, independent of any enumeration of the partial recursive functions. Moreover, it is of the same general character as definition by recursion in the wider set-theoretical setting.

#### 4.3. Partial Recursive Functionals of Finite Type over the Natural Numbers

Notions of primitive, general and partial recursive functionals have been extended to various *finite type (f.t.) structures*  $\mathcal{M} = \langle M_\tau \rangle$  over the natural numbers, where  $M_0 = N$  and  $M_{(\tau,\sigma)}$  consists of certain (possibly partial) operations  $f$  from  $M_\tau$  to  $M_\sigma$ . In set-theoretical terms one can define the structure HTF of *hereditarily total functionals of f.t.* by  $M_{(\tau,\sigma)}$  = the set of all (total)  $f: M_\tau \rightarrow M_\sigma$ . Gödel in 1958 (cf the 1990 collection) introduced a notion of

primitive recursive functional which makes sense for objects in this type structure (but also for much narrower structures). Kleene 1959 dealt with partial recursive functionals  $F(f_1^{(n_1)}, \dots, f_m^{(n_m)}) \simeq y$  where  $f^{(n)}$  is an object of pure type  $n$ , i.e., belongs to  $M_n$ , where those are defined by  $M_{n+1} = M_{(n,0)}$ . The type structure of hereditarily total functions can be reduced to that of pure types. There are also reasonable extensions of the notion of partial recursive functional to suitable structures of hereditarily continuous functions of f.t. (Kleene 1959a, Kreisel 1959) and hereditarily monotonic partial functions of f.t. (Platek 1966 and § 5.3. below). It is not possible to explain all these notions without going into considerably more detail than we have space for here.<sup>5</sup> However, this should begin to suggest that the idea of partial recursive function(al) makes sense for a variety of structures  $\mathcal{M}$  of objects which are not, themselves, effectively given. This constitutes the next chapter in the relativization of recursion theory.

## 5. Generalized Recursion Theory

### 5.1. Background and Overview

The development of recursion theory discussed up to this point has taken us from the early 1930's up to the late 1950's. The first part of this development, which took place roughly up to the mid-'40's, was devoted to foundational work, applications, and systematic organization of the subject. In that period, except for the introduction of Turing's concept of computation relative to an "oracle" which remained untouched until 1944, recursion theory was conceived of in *absolute* terms. The second period of development sees a branching of the subject into distinctive subfields and with increasing specialization and technical sophistication. Only two of these have been discussed above, namely *degree theory* and *the theory of recursive functionals*. A third area, which we shall not attempt to describe, concerns what is sometimes called *hierarchy theory*, and in particular the extension of the arithmetical hierarchy to the *hyperarithmetical hierarchy* through all *recursive ordinals* (i.e., ordinals with a recursive order type); for this cf. Hinman 1978 and Sacks 1990. All of these concern *relativization* in one way or another, the last by transfinite iteration of certain "jump" operations on sets. At the same time, the initial motivation to secure an "absolute" concept of effective computability in order to establish the effective unsolvability of classical problems still maintained its force,

---

<sup>5</sup> See Odifreddi 1989, pp. 199–201, for a survey of various notions of recursion over higher-type structures.

with the eventual resolution (as has already been noted) of such outstanding questions as the word problem for groups and Hilbert's 10th problem. Still another use of both the basic theory and the theory of recursive functionals was in a more positive direction, namely to supply recursion-theoretic semantics for intuitionistic formal systems, by means of so-called *recursive realizability interpretations* (for this cf. the survey article Troelstra 1977).

During this same over-all period from the 1930's up to the late 1950's, mathematical logic as a whole was undergoing considerable development, following roughly the same pattern: in the pre-war period, by means of foundational and organizational work with basic applications, and in the post-war period through a split-up into more specialized and technically sophisticated research programs. But the field as a whole tended to be compartmentalized, with little interaction between the different directions of work into what are still regarded as the main branches of logic: *set theory*, *model theory*, *recursion theory* and *proof theory*.

One landmark event signaled the breakdown of this compartmentalization, namely a six-week long Institute in Symbolic Logic held in the summer of 1957 at Cornell University. This brought together leading research workers and their students from all the different fields of logic and encouraged a process of intercommunication and interaction which has continued unabatedly ever since. While each of the branches of mathematical logic still maintains a distinctive character and body of concerns, it is difficult to work in any one of them nowadays without using knowledge and techniques from one or more of the other branches.

Perhaps more than any of these, recursion theory was to see an infusion of concepts, methods and examples from all the other branches which significantly affected its development in the following years. This was to transform its conceptual arena from the natural numbers (and related effectively enumerated structures such as word systems) in what has come to be called *ordinary recursion theory* (o.r.t.), to quite *general structures*, thus opening up the development of what is called *generalized recursion theory* (g.r.t.). This in turn has followed two lines: (i) generalization of recursion theory to various *structures of sets and ordinals*, and (ii) generalization of recursion theory to (more or less) *arbitrary* ("abstract") *structures*.

The subject of g.r.t. is much more difficult to describe than the material discussed up to now in § 2–4, because of the welter of conceptual approaches, structures to which they are applied, and results obtained. Some initial impression of this variety can be obtained from the books Barwise 1975, Fenstad 1980, the conference volumes Fenstad and Hinman 1974, Fenstad, Gandy, and Sacks 1978, and the survey articles by Shore, Kechris and Moschovakis,



Aczel, and Martin (in that order) in the *Handbook of Mathematical Logic* (Barwise 1977). Last, but not least, to be mentioned is the survey and critical assessment of g.r.t. by Kreisel (1971) which represents the situation mid-stream. The very recent (but long-awaited) book by Sacks (1990) now fills in much of the technical picture for recursion theory on sets and ordinals, though not for g.r.t. on arbitrary structures.

Given the complexity and heterogeneity of these developments, nothing short of a full survey and new critical assessment would do justice to g.r.t. My purpose in the following is just to give the reader *some* sense of how *some parts* of this proceeded, as part of a picture of the further relativization of recursion theory, and in particular to emphasize the conceptual shift to a structural view of its subject matter.

## 5.2. Computability over Sets and Ordinals

Here I follow in part the survey by Shore 1977, which gives a good introduction with historical background together with references that I shall not repeat (cf. also Kreisel 1971).

A notion of recursive function of ordinals was introduced by Takeuti in 1960 by means of schemata, where the schema for primitive recursion is expanded to all ordinals by taking *sup* at limit ordinals  $x$ ,

$$(1) f(x) = \sup \{ g(y) \mid y < x \}.$$

Another generalization of recursion theory to ordinals was provided by Machover in 1961 using an extension of the Herbrand-Gödel equation calculus with certain infinitary rules of inference, and by Lèvy in 1963 using an analogue of Turing machines; both Machover and Lèvy observed that one could work just as well with the ordinals less than a regular cardinal, since that is closed under suprema (1). A still further refinement was made by Kripke in 1964 and Platek in 1966, who realized that a much wider class of ordinals, called the *admissible ordinals*, support a reasonable generalization of recursion theory. Kripke again used a form of the equation calculus, while Platek used both a definition by schemata and one by generalized computers. As described in Barwise 1975, p. 3, an ordinal  $\alpha$  is called admissible if for every  $\alpha$ -(partial) recursive function  $f$  of ordinals, whenever  $x < \alpha$  and  $f(x)$  is defined, then  $f(x) < \alpha$ , where, moreover,  $f$  is  $\alpha$ -(partial) recursive if its values can be computed by an "idealized computer capable of performing computation of less than  $\alpha$  steps".

Meanwhile (1963–65), Kreisel and Sacks had been developing recursion theory on the ordinal  $\alpha = \omega_1^{\text{CK}}$ , the least non-recursive ordinal in the sense of

Church and Kleene, (the recursive analogue of the least uncountable ordinal  $\omega_1$ ); this turns out to be the first admissible ordinal greater than  $\omega$  (the ordinal of the natural number structure  $(N, <)$ ). Sacks sought to generalize results of degree theory from o.r.t. to recursion theory on  $\omega_1^{CK}$ . For this it was necessary to have a suitable generalization of the relation  $A \leq_T B$  of relative computability for arbitrary sets  $A, B$ . The crucial ingredient was supplied by Kreisel in the form of a generalized notion of finiteness. Extended directly to arbitrary admissible ordinals  $\alpha$  and subsets  $A$  of  $\alpha$ , his proposal was to define:

- (2)  $A$  is  $\alpha$ -finite if  $A$  is  $\alpha$ -recursive and bounded, i.e., if  $A$  is contained in  $\beta$  for some  $\beta < \alpha$ .

Now the Kreisel-Sacks definition of  $A \leq_\alpha B$  (Turing reducibility generalized to recursion theory on an admissible ordinal  $\alpha$ ) is, roughly speaking, that every  $\alpha$ -finite subset of  $A$  can be determined in an  $\alpha$ -effective way from some  $\alpha$ -finite subsets of  $B$  and its complement (to  $\alpha$ ), and similarly for every  $\alpha$ -finite subset of the complement of  $A$ . Before long, Sacks and his students were extending one result after another from degree theory in o.r.t. to arbitrary admissible ordinals. In particular, in 1972, Sacks and Simpson established the analogue of the Friedberg-Muchnik solution to Post's problem: there exist  $\alpha$ -r.e. sets  $A, B$  such that neither  $A \leq_\alpha B$  nor  $B \leq_\alpha A$ . This makes use of an extension of the priority method to admissible ordinals, for which a full technical exposition is now to be found in *Sacks 1990*.

Recursion theory on admissible ordinals also gave rise to a recursion theory on sets via the intimate relation between ordinals and constructible sets, in the sense of Gödel (cf. 1940 in the collection 1990)). Another form of recursion on sets, called *E-recursion*, and distinct from admissible recursion theory, was introduced (independently) by Normann and Moschovakis around 1978. A number of generalizations of degree theory have also been obtained for *E*-closed sets (also given an exposition in *Sacks 1990*). Though there were many motivations for generalization of recursion theory to ordinals and sets, and these have been satisfied to a large extent by the subsequent developments, the research program of generalized degree theory has been that direction which has been pursued most vigorously, and again with the most technically difficult results.

### 5.3. Computability over General Structures

The idea of generalizing recursion theory to (more or less) arbitrary structures also began early in the 1960's. The article *Kreisel 1971* provides a comprehensive source for the developments up to that publication, and references cited

there will not be repeated here. One of the first proposals was made by Fraïssé in 1961, in model-theoretic terms. A variety of subsequent proposals receiving special attention (among others) were those due to Lacombe, Montague, Moschovakis, Platek, and Friedman; we shall sketch only the latter two approaches here, in reverse historical order.

The work of Friedman (1971) generalized notions of Turing machines and register machines to arbitrary first-order structures  $\mathcal{M} = \langle M_0, R_1, \dots, R_l, g_1, \dots, g_m \rangle$ , where now  $R_j$  are relations and the  $g_j$  may be partial functions. If the relation  $x = y$  on  $M_0$  is to be counted as computable it must be included as one of the basic relations, but that is not assumed in general. In Friedman's generalization of register computability to an arbitrary structure  $\mathcal{M}$ , each register is empty or contains an element  $a$  of  $M_0$ . Then the actions specified by the instructions are of the form

- (1)  $r_i := g_j(r_{n_1}, r_{n_2}, \dots)$  and
- (2) if  $R_j(r_{n_1}, r_{n_2}, \dots)$  then go to  $I_k$  else to  $I_l$ .

The meaning of (1) is to replace the contents of  $r_i$  by  $g_j(a_1, a_2, \dots)$  where  $a_l$  is in the  $n_l$ 'th register, and of (2) is to perform a conditional transfer with test  $R_j(a_1, a_2, \dots)$ . Then a partial  $f$  is *register computable* over  $\mathcal{M}$  if  $f(a) = b$  just in case a computation with  $r_1 = a$  as input terminates with  $r_0 = b$  as output. (And similarly for  $n$ -ary  $f$ ). This generalizes the Shepherdson and Sturgis notion by taking  $\mathcal{M}$  to be the structure  $\langle N, R_1, 0, sc, pd \rangle$  where  $R_1$  is the unary relation  $\{0\}$ , i.e.,  $R_1(x)$  iff  $(x = 0)$ ; equivalently we may take  $\mathcal{M}$  to be the structure  $\mathcal{N} = \langle N, =_N, 0, sc, pd \rangle$ . It also yields the relation  $f \leq_T g$  as a special case, since that holds just in case  $f$  is computable over  $\langle N, =_N, 0, sc, pd, g \rangle$ . Friedman's notion generalizes directly to many-sorted structures. Then he defines  $f$  to be *register computable over  $\mathcal{M}$  with counting* if it is computable on the combined structure  $(\mathcal{M}, \mathcal{N})$ . Friedman also generalized computability by Turing machines to arbitrary  $\mathcal{M}$ , where the contents of a Turing tape cell may be empty or filled by an element of  $\mathcal{M}$ .

Kreisel (1971, p. 144) asked whether there is such a thing as 'an extension of Church's thesis to general (abstract) structures'. In his discussion of this (op. cit., pp. 175 ff.) he points out that "Evidently two elements are involved in Turing's analysis, ... the objects on which we operate, [and] the instructions or rules of computation". According to him, Turing's analysis requires a restriction on how the objects of computation may be presented to us, and what operations on them may be assumed. From this point of view, Turing computability on the structure  $\langle N, \dots, g \rangle$ , where  $g$  is not recursive, is not a suitable structure for computation. Nor would a structure  $\langle N, N_1, \dots \rangle$  be admitted, for  $N_1 = (N \rightarrow N) = \{g \mid (g : N \rightarrow N)\}$ . Clearly, the essence of

Friedman's generalization of register computability is to give up any restriction on how the objects on which we operate are presented to us, but to maintain the form of the instructions or rules of computation. *Shepherdson 1988* has extended Gandy's principles for mechanisms to arbitrary structures, and it is there argued that Friedman's "machines" lead to a general form of the Church-Turing Thesis. As has already been stated, we do not take issue here with such a position, one way or another.

However, it should be noted that not all structures for which we have a reasonable generalization of recursion theory fall under Friedman's definitions. In particular, recursion theory on admissible ordinals or sets beyond the natural numbers don't come out as special cases; the reason is that they embody essentially *infinitary operations*, such as sup.

This brings us to Platek's generalization of recursion theory to arbitrary  $\mathcal{M}$ , carried out in his (unpublished) dissertation 1966. Besides explicit functional definition (using the operations and relations of  $\mathcal{M}$  as basic), this takes the Recursion Theorem

$$(3) f = F(f)$$

as its central means of definition. In order to make sense of this as providing the *least fixed point*  $f = FP(F)$ , it must be assumed at the minimum that  $F$  is a monotonic functional. But then the question arises, which  $F$ , specifically, are to be used? Platek's answer was that they must in turn be generated by the recursion theory over  $\mathcal{M}$ . He thus introduced a type structure HMF of *hereditarily monotonic functionals over*  $\mathcal{M} = \langle M_0, \dots \rangle$ . For this, a relation of inclusion is defined at each type, with  $f$  contained in  $g$  at type  $(\sigma, \tau)$  if for all  $x$  in  $M_\sigma$ ,  $f(x)$  is contained in  $g(x)$  at type  $\tau$ . Then  $M_{(\sigma, \tau)}$  is taken to consist of all monotonic  $f: M_\sigma \rightarrow M_\tau$  in this inclusion relation. Now for each  $F$  in  $M_{(\rho, \rho)}$  where  $\rho = (\sigma, \tau)$  there is a least  $f$  in  $M_\rho$  satisfying the fixed-point equation (3). Finally, the operation  $FP_\rho: M_{(\rho, \rho)} \rightarrow M_\rho$  is itself in  $M_{((\rho, \rho), \rho)}$ . With each collection  $\mathcal{F}$  of functionals in this type structure over  $\mathcal{M}$  is associated the collection  $Rec(\mathcal{F})$  generated by explicit definition and all the fixed-point operators  $FP_\rho$ ; the basic operations and relations of  $\mathcal{M}$  are built into  $\mathcal{F}$ .

The type structure HTF( $M_0$ ) of hereditarily *total* functionals over  $M_0$  can be extracted from HMF( $M_0$ ). In particular, when  $M_0 = N$ , Platek recaptured Kleene's 1959 partial recursive functionals over  $N$  by taking  $\mathcal{F} = \{0, sc, pd\}$ , and Kleene's notion of (higher type) recursion in some particular functions or functionals  $F_1, \dots, F_m$ , by taking  $\mathcal{F} = \{0, sc, pd, F_1, \dots, F_m\}$ . In this way we can incorporate infinitary operations, e.g., the functional  ${}^2E$  of Kleene, with

$$(4) {}^2E(f) = 0 \text{ if } (\exists x) (f(x) = 0) \text{ and } {}^2E(f) = 1 \text{ otherwise, for } f: N \rightarrow N.$$

(An important result of *Kleene 1959* was that the functions recursive in  ${}^2E$  are exactly the hyperarithmetical functions). But Platek also recaptured recursion on admissible ordinals and sets, by taking recursion in the functional *Sup* given by

$$(5) \text{Sup}(f, x) \simeq \sup \{f(y) \mid y < x\}.$$

Is it necessary to go through all higher types in order to find the functions of type 1 in  $\text{Rec}(\mathcal{F})$ ? One of the principal results of *Platek 1966* is that if every member of  $\mathcal{F}$  is of type level  $\leq 2$  and  $f$  is in  $\text{Rec}(\mathcal{F})$  and of level  $\leq 2$ , then we need only use the schemata for explicit definition and FP applied in type levels  $\leq 2$  in order to obtain  $f$ . The above examples with  ${}^2E$  and *Sup* meet these conditions.

While Platek's approach is of impressive generality and builds on a natural basic idea (recursion as given by the FP operator), it does not cover all the cases one would want to include. In my paper *Feferman 1977*, pp. 376–7, I discussed certain limitations of Platek's theory. In brief, these are:

- (i) It is assumed that there are pairing and projection functions on  $\mathcal{M}$  in the basic  $\mathcal{F}$ , as well as distinct elements 0, 1 from  $M$ ; thus  $M_0$  contains an image of the natural numbers and the possibility of enumeration. It is preferable to separate out the natural numbers by a different basic sort if they are to be used at all.
- (ii) The theory does not generalize relational notions of computability, for which the paradigm is the Post-Smullyan approach (cf. *Smullyan 1961*, *Fitting 1987*).
- (iii) Details of the extraction of the HTF type structure from the HMF type structure are very messy for types  $>2$ , and this makes extraction of the general *Kleene 1959* notions very complicated.

Moving beyond the preceding, in the mid-1970's Moschovakis and I independently proposed to get around such defects by treating recursion in higher types as a special case of recursion on arbitrary structures, rather than as the means to define it. As I put it (1977, p. 373): "In contrast to Platek, higher-type structures are regarded here as just further examples which are to be *subjects* of the notions of g.r.t. rather than *tools* to explain the notions". My approach was sketched in *Feferman 1977*, but all the detailed work has been carried out by Moschovakis, first in collaboration with Kechris (*Kechris and Moschovakis 1977*) for the special case of Kleene recursion in higher types, and then more generally in *Moschovakis 1984*, among other publications. Basically, the notions concern type level 2 functionals  $F$  with arguments chosen from a collection  $\mathcal{R}$  of type level 1 relations over a many-sorted structure

$\mathcal{M} = \langle \langle M_k \rangle, \dots \rangle$ , where  $\mathcal{R}$  is closed under unions of chains. In particular,  $\mathcal{R}$  may be chosen to be all (sorted) partial functions over  $\mathcal{M}$ , or all (sorted) relations over  $\mathcal{M}$ . Then for any collection  $\mathcal{F}$  of type level 2 monotonic functionals  $F$  over  $\mathcal{R}$ , one defines  $\text{Rec}(\mathcal{F})$  to be the least collection of objects of type levels  $\leq 2$  generated by explicit definition and the least fixed point operators  $FP$ . This gives rise to a theory of great generality, although it is still not clear whether it covers *all* cases for which we have a reasonable generalization of recursion theory.<sup>6</sup>

There is one final significant step that Moschovakis has taken in his 1984 paper. This is to consider functionals  $F$  *operating across structures*. That is, given a class  $\mathcal{K}$  of structures of the same similarity type, one can give meaning to objects  $F$  defined over  $\mathcal{K}$  such that for each  $\mathcal{M}$  in  $\mathcal{K}$ ,  $F(\mathcal{M})$  is a functional  $F_{\mathcal{M}}$  over  $\mathcal{M}$ , such that the functionals  $F_{\mathcal{M}}$  all *act in the same way*. In other words, this provides a notion of *uniform computability across structures*.<sup>7</sup> The significance of this for actual computability will emerge in the next section.

## 6. The Role of Notions of Relative Computability in Actual Computation

### 6.1. Computational Practice and the Theory of Computation

The kinds of mechanisms we have in mind here are high-speed, digital, general-purpose computers, from PCs to mainframes. For these, the aim of computational practice is to produce hardware and software that is reliable, efficient, flexible, versatile and user-friendly. The aim of the theory of computation is to aid engineers in the design of hardware and software meeting these requirements, by providing a body of concepts around which to organize experience and a body of results predicting correctness, efficiency, and versatility. Theory also serves to set limits to what is feasible, and thus provides warning

<sup>6</sup> In particular, Part II of the paper *Feferman 1977* was concerned with the question whether the notions of partial recursive functional of *hereditarily total continuous objects* ("countable functionals") developed by Kleene and Kreisel in 1959 is covered by this theory. As far as I know this is still open. However, a related and more extensive notion of recursion on the hereditarily *partial* continuous functionals was shown to be accounted for by recursive schemata in the above sense.

<sup>7</sup> Something like this was anticipated in a lecture for the Association for Symbolic Logic that I delivered in 1969 (cf. JSL 35 (1970), p. 179); regrettably, the material was never published, though I circulated handwritten notes, "Uniform inductive definitions and generalized recursion theory," at the time. Cf. also *Kreisel 1971*, pp. 147–8.

signals for when these limits are approached. The theory of computation employs logic and mathematics ranging from the most concrete combinatorial kind to the most abstract, algebraic, and topological kind. The following tells some "stories" about the role notions of relative computability play in computational practice, that sits somewhere in the middle between the two extremes of the theory of computation. Unlike the preceding sections of the paper it is neither historical nor (for the most part) does it concern specific results.

The literature in theoretical computer science accepts the Church-Turing Thesis in principle, of course with the proviso that this must be supplemented by an assessment of time and space requirements for practice. Sometimes it is said that the notion of finite automaton must be substituted for that of Turing machine (or equivalent), to reflect the actual limitations on space. However, in practice memory (storage) is expandable, and the automaton model does not account for that. On the other hand, it is generally recognized that Turing machines themselves do not provide a realistic model of actual mechanisms since "... they are confined to specific data structures (the tapes) which have artificially high large access time (because in order to read a bit far away on a tape the respective head has to travel over all cells in between)". (*Maass and Slaman 1989*, p. 80). Register machines instead provide a more realistic model of random access memory in practice (cf. loc. cit. and *Aho, Hopcroft, and Ullman 1974*). Moreover, at least one style of programming is directly keyed to that mode, namely ("von Neumann") imperative-style, with assignment statements, e.g., as in PASCAL. But the theory of computation must account for a number of other programming styles such as functional programming or logic programming, which are less directly related to the nature of the hardware. For all of these and even for imperative style programming, the details of the underlying mechanism are largely considered to be irrelevant. Thus the question arises what the significance of the Church-Turing Thesis is for computational practice; on the face of it, the thesis seems to be a matter of basic creed which has nothing to do with day-to-day computational life. A four-fold response was provided to me by P. Odifreddi in correspondence, summarizing points in his introduction to the collection "Logic and Computer Science" of which he is the editor (1990): (i) The notion of universal Turing machine is the idealization (and conceptual precursor) of modern *all-purpose computers*; (ii) the Enumeration Theorem shows the *equivalence of programs and data*, basic for stored program machines; (iii) The proof of Kleene's Normal Form Theorem for partial recursive functions provides the theory underlying *interpreters*; and (iv) various definitions of recursiveness provide the *computational core (and style) of different pro-*

*gramming languages* (cf., e.g. PASCAL, PROLOG and LISP). In personal conversation, (and forthcoming work), W. Sieg has further emphasized the constraints placed on actual computation by the theoretical analysis stemming from Turing (1936–7) and leading to Gandy's very general principles for mechanisms (1980): after all, what counts as computations in its every-day sense can't be completely arbitrary.

While I can hardly disagree with these points (and have already brought out (i) and (ii) in § 2), it will be argued here that nevertheless, notions of *relative* computability have a much greater significance for practice than those of *absolute* computability. The reason is very simple: as with all forms of technology, the requirements of efficiency, reliability, and usability force an organization of the devices and their control into *conceptual levels* and at each level into interconnected *components*. At the hardware level, one has a breakdown into such gross components as a central processing unit (CPU), memory locations, both read-only (ROM) and random access (RAM), a clock, etc.; then for each of these one has further refinements into subcomponents such as adders, down finally to the level of individual switches. At each level one depends on standard designs but always subject to improvements, so that if any one component is changed, the performance of the other components is not affected. Moreover, if the whole material basis of the technology is changed from, say, chips to fiber optics, the organization of components need not be modified. Nor is there a simple dichotomy between the hardware-software levels (or trichotomy, if one adds in the user). Rather, there is a step-wise ascent from hardware to software or, from the point of view of the programmer, descent from the programming language through a compiler or interpreter down to assembly language and, finally, to "machine" language. And for the programmer there are, to begin with, the shifts in level from informally stated problems and tasks to their mathematical or symbolic formulation, down to a concrete program in one language or another, all well-captured in the slogan of "top-down design" (cf., e.g., *Alagic and Arbib 1978* and *Harel 1987*).

While notions of relative computability have some connection with the different conceptual levels of organizations in hardware and software, what is rather emphasized in the following is their significance, at a given level, for *modular organization*, i.e., for *how things are packaged*, and *how they fit together*.

## 6.2. Built-in Functions and Black Boxes

To become more specific, let us return to Turing's "oracle" machines and the relation  $f \leq_T g$ . Actual computers have a variety of built-in functions  $g$ , whose



values may be called at any point in a program. These are for arithmetical operations on integers such as  $+$ ,  $-$ ,  $*$ , *quot*, *rem*; Boolean operations such as 'and', 'or', 'not'; operations from integers to Booleans such as *lesseq*; and sometimes also operations from and to (approximate) real numbers such as *sqrt*, *sin*, *log*, etc. As far as the programmer is concerned, each of these is given by a "black box" – which is just another name for an "oracle" – and a program to compute a function  $f$  from one or more of these  $g_1, \dots, g_m$  is really an algorithm for computation of  $f$  relative to  $g_1, \dots, g_m$ . Such an algorithm can thus build in commands to apply one of the  $g_j$  to arguments which arise in the course of the computation. Moreover, for certain purposes, measures of complexity can also be relativized to the black boxes, e.g., they might be assigned unit cost or even 0 cost.<sup>8</sup>

### 6.3. Functional Aspects of Programming

These are both *implicit* and *explicit*. Examples of the former are provided by *flowchart analyses* for certain programming languages. Consider programs  $\Pi$  for a register machine. At any point in a computation the operation of such  $\Pi$  is determined by the *state*  $s$  of the contents of the various registers and the effect of  $\Pi$  is to change  $s$  to  $\Pi(s)$ . Thus  $\Pi$  may be considered as (determining) a function  $\Pi: S \rightarrow S$  where  $S$  is the set of all states. Now in a fragment of a flowchart program,

$$(1) \rightarrow \Pi_1 \rightarrow \Pi_2 \rightarrow$$

indicates the composition  $C(\Pi_1, \Pi_2)$  which has the effect

$$(2) C(\Pi_1, \Pi_2)(s) = \Pi_2(\Pi_1(s)).$$

The construction  $C$  here may be considered to be a functional on  $S^S \times S^S$  where  $S^S = \{\Pi \mid \Pi: S \rightarrow S\}$ . As another example, conditional branching, whose flowchart follows  $\Pi_1$  if  $R$  is true and otherwise  $\Pi_2$ , where  $R$  is contained in  $S$ , gives rise to the functional

$$(3) B_R(\Pi_1, \Pi_2)(s) = (\text{if } s \text{ is in } R \text{ then } \Pi_1(s), \text{ else } \Pi_2(s)).$$

Similarly, we may treat such program constructions as

$$(4) \text{ while } R \text{ do } \Pi, \text{ and}$$

$$(5) \text{ do } \Pi \text{ until } R,$$

---

<sup>8</sup> Several people have suggested to me that *interactive computation* exemplifies Turing's "oracle" in practice. While I agree that the comparison is apt, I don't see how to state the relationship more precisely.

as functionals of  $\Pi$  and  $R$  (or of a program for  $R$  as a function from  $S$  to  $\{T, F\}$ ). In these cases, the functionals may yield partial functions on states to states, as values.

Examples of *explicit* functional operations are given by functional-style programming languages such as LISP and ML. In these we can form expressions involving functional recursion such as  $F$  defined by the following:

$$(6) F(g, h; x) = [g(0) \text{ if } x = 0, \text{ else } h(F(g, h; x - 1), g(x))],$$

which has the solution  $f = F(g, h)$  with

$$(7) f(0) = g(0),$$

$$f(x') = h(f(x), g(x)).$$

For example,  $F(g, +, x)$  and  $F(g, *, x)$  yield, respectively, the sum and product of terms  $g(y)$  for  $y \leq x$ .

The use of higher-order functions permeates functional programming languages, cf. the text *Reade 1989*. They are generally based on some form of the untyped  $\lambda$ -calculus, though flexible ("polymorphic") systems of typing have also been imported (cf. op. cit. as well as *Feferman 1990* for references). In these languages, programs are represented by expressions, and operations on programs such as composition, conditional branching, iteration, etc. are represented by compound expressions. In the strictly typed  $\lambda$ -calculus there are rigid rules that govern the compounding of expressions, and thus tell exactly how the corresponding programs may be interconnected. In untyped calculi with polymorphic type assignment systems (cf., e.g., *Mitchell and Harper 1988*) such rules are considerably more flexible, permitting combinations forbidden in the strictly typed calculus but still providing for sensible interconnections of the corresponding programs. In terms of the theme of § 6.1 above, these give systematic ways of representing the modular construction of software.

Research on type systems, logics and semantics for functional programming languages is still being carried on vigorously by a number of authors (cf. the works cited above for further references).

#### 6.4. Abstract Data Types.

All programming languages deal with types of expressions either internally, within the syntax, or externally, in the semantics. One generally has such basic data types as integers, Booleans, and reals, and then general type constructions, such as for lists, arrays, stacks, queues, sets, trees, streams, etc. In functional programming languages the concern is also with higher-order data

types such as for functions and functionals. Again, there are many approaches to dealing with these concepts, and research is ongoing. The purpose here is to show how certain ways of looking at these connect with relativized recursion theory, and more specifically with recursion theory over general structures.

From a semantical or external point of view, abstract data types are either specific structures  $\mathcal{M}$  considered independently of the form of representation of their elements, in other words as *isomorphism types*, or as *collections*  $\mathcal{K}$  of *structures* of a given similarity type. In either case, these structures may be prescribed by *axiomatic defining conditions*  $A$ , e.g., by equations or Horn clauses. In general such axiom systems are not categorical, unless supplemented by some second-order conditions, e.g., that  $\mathcal{M}$  is the least structure satisfying  $A$  (or is an *initial structure* for  $A$ ), or that  $\mathcal{K}$  consists of all finite structures satisfying  $A$ , etc. In whatever way  $\mathcal{M}$ , resp.  $\mathcal{K}$ , are prescribed, one can give a semantics for programs on these structures using one of the generalizations of recursion theory mentioned or described in 5.3. For example, *Tucker and Zucker 1988* consider various forms of schematic definability, while *Moschovakis 1984* uses the general form of inductive defining schemata. An interesting result from the latter (cf. p. 326) is that uniform global recursion on the class of finite structures with a linear ordering gives exactly the polynomial time computable relational queries for these structures (for which notion, cf. *Chandra and Harel 1980*).

Uniform global recursion provides a much more realistic picture of computing over finite data structures than the absolute computability picture, for finite data bases are constantly being updated. As examples, we may consider computations on weather data (given by finite samples from a continuous space) for weather prediction, or the status of communication lines for routing in a telephone system, or airline reservation systems, and so on, with endless practical examples.

Limitation of space here does not allow me to go into internal or syntactic representation of abstract data types. For some approaches cf. *Mitchell and Plotkin 1986* and *Feferman 1990*.

### 6.5. Degrees of Complexity

So far we have only considered the significance for computation theory of notions of relative computability other than those from the theory of degrees of unsolvability. But the latter would seem to provide a *prima facie* case not only of application of notions, here to *complexity theory*, but also of methods and results. There is some dispute, though, as to whether the latter subject should be considered a part of computation theory or a part of recursion

theory. Be that as it may, let us consider the situation as it appears at present. Here we rely on such sources as the venerable *Garey and Johnson 1971* and the up-to-date exposition *Balcázar, et al 1988*.

The theoretical basis for predicting the relative efficiency of algorithms lies in the assignment of time and space measures of complexity to algorithms. For example, one bounds how long it takes to compute a function  $f(n)$  by a given algorithm, as a function of the size  $|n|$  of input in binary notation. The crudest distinction puts tractable problems in the class that can be computed in  $O(p(|n|))$  time for some algorithm and polynomial  $p$ , and intractable problems in the class that require at least  $O(2^{|n|})$  time for any algorithm. A function computable by an algorithm of the former kind is said to be *polynomial-time computable*, and the class of these is denoted  $P$ . A set  $A$  of natural numbers (or the decision problem for membership in  $A$ ) is in class  $P$  if  $c_A$  is in  $P$ . There are a number of decision problems that are not obviously in class  $P$  but lie in a class that is not as complicated as those requiring exponential time. This is denoted  $NP$ , for *non-deterministic polynomial time computability*. Roughly speaking, these are problems  $A$  for which one can check of a given  $n$  that  $n$  is in  $A$ , when in fact it does belong to  $A$ , by means of some certifying evidence which is itself verified to be such in  $P$ -time. For example, the problem of satisfiability of a formula in the classical propositional calculus is in class  $NP$ . On its face, it is hard to decide whether such a formula is satisfiable since we must set up a truth-table for it, and if it contains  $n$  literals, that contains  $2^n$  lines, each of which must be checked. But if a truth assignment  $s$  actually is one which makes the formula satisfiable, that fact can be checked in polynomial time. For problems  $A$  in the natural numbers the notion of  $NP$ -computability can be formulated as definability in the form

$$(1) x \text{ in } A \text{ if and only if } (\exists y) [ |y| \leq p(|x|) \ \& \ R(x,y) ],$$

where  $p$  is a polynomial and  $R$  is in the class  $P$ .

The non-deterministic aspect of such lies in the fact that one may not have a feasible method of choosing in advance, given  $x$  in  $A$ , a  $y$  which quickly certifies that  $x$  is in  $A$ . The form (1) readily suggests an analogy

- (2)  $NP \sim$  recursively enumerable, along with
- (3)  $P \sim$  recursive.

Moreover, one has a concept of *polynomial-time reduction* of problems which is analogous to that of Turing reducibility, where  $A \leq_p B$  holds if, roughly speaking, there is an algorithm which would transform any  $P$ -algorithm for  $B$  into one for  $A$ . Thus

$$(4) (\leq_p) \sim (\leq_T).$$

This immediately suggests the notion of *NP*-completeness analogous to that of Turing completeness in the class of r.e. sets; *B* is called *NP*-complete if every *NP* set *A* is *P*-reducible to *B*. There are, indeed, *NP*-complete problems; this was the major result of *Cook 1971*, who showed that the satisfiability problem for the propositional calculus is one such. Since then, a number of other problems that arise naturally in practice have been shown to be *NP*-complete, including the "travelling-salesman" problem and the Hamiltonian path problem (cf. *Garey and Johnson 1971*). So far, so good: the parallels to degree theory are persuasive. One now comes to posing the analogue of the Church-Turing existence of effectively unsolvable problems:

(5) Does  $P = NP$  ?

No answer is yet known to this, though it is generally conjectured that  $P \neq NP$ . But here one has a breakdown in the analogy. Namely, effective unsolvability in ordinary recursion theory relativizes, but the  $P = NP$  question does not. That is, the halting problem *H* (or diagonal halting problem *K*) which demonstrates the existence of r.e. but non-recursive sets is such that, relative to any set *A*, we have

(6)  $H^A$  is not  $\leq_T A$ .

Put in other terms, for any *A*,  $Rec^A$  is properly included in  $R.E.^A$ . However, *Baker, Gill, and Solovay 1975* proved that

(7) there exist *A* such that  $P^A = NP^A$ , though there exist *B* such that  $P^B \neq NP^B$ .

If, as is generally conjectured,  $P \neq NP$ , it would be natural to further investigate the analogue to Post's problem:

(8) Do there exist *A* in *NP* which are not in *P* and not *NP* complete?

Here there is a positive answer (assuming  $P \neq NP$ ) due to Ladner in 1975; cf. *Balcázar, et al 1988*, p. 156. But it seems that none of the non-trivial techniques or results of degree theory has so far been of any use for this or any other results in complexity theory. Naturally, future developments may change that situation.

Finally, we should mention the development of *hierarchies* similar to the arithmetical, e.g., the *polynomial-time hierarchy* introduced by Meyer and Stockmeyer in 1973 and treated in *Balcázar, et al 1988*, ch. 8. However, many of the basic questions about this are open, such as whether the entire hierarchy goes beyond *P* or, alternatively, simply collapses.

## 6.6 Conclusion

Our final section has explored the question of the relevance of the mathematical theory of computability – in the guise of recursion theory – to the theory of computation, insofar as *that* is supposed to be a theory of computational practice. To avoid misunderstanding, I do not believe recursion theory is to be valued only if it can have such applications. Aside from the clear philosophical value of having a fundamental analysis of the notion of computation, there have been plenty of applications in logic and mathematics to justify its existence, if justification by external relevance is called for at all. But even in the most rarified and recondite parts of the subject such as degree theory, persuasive intrinsic reasons can be given for the lines of development that have been taken and for the continued pursuit of internally driven problems (along with the feeling that '... we have to do it because it's *there*'). Be that as it may, the case has been made here that while notions of relativized (as compared to absolute) computability theory are essentially involved in actual hardware and software design, the bulk of methods and results of recursion theory have so far proved to be irrelevant to practice. Whether and how these disciplines might be brought closer together remains for the future to tell.

To conclude on a more positive but more speculative note, I think it can be argued that – whatever a "fundamental" theory may tell us is the basic or underlying mechanism for given technological processes or systems – it is necessary for our design and use of such to think of them at various conceptual levels and with various modular forms of organization, that is we *must think of them in structural terms*. The story of Turing's "oracle" and its significance for actual computability is but one example among many of this characteristic *modus operandi* of human intelligence.

## References

- A. Aho / J. E. Hopcroft / J. Ullman (1974), *The design and analysis of computer algorithms*, Addison-Wesley.
- S. Alagic / M. A. Arbib (1978), *The design of well-structured and correct programs*, Springer-Verlag.
- T. P. Baker / J. Gill / R. Solovay (1975), "Relativizations of the  $P = NP$  question", *SIAM J. Comp.* 4, 431–442.
- J. L. Balcázar / J. Díaz / J. Gabauró (1988), *Structural complexity I*, Springer-Verlag.
- J. Barwise (1975), *Admissible sets and structures*, Springer-Verlag.
- J. Barwise (1977) (ed.), *Handbook of mathematical logic*, North-Holland.
- A. K. Chandra / D. Harel (1980), "Computable queries for relational data bases", *J. Comp. Syst. Sci.* 21, 156–178.

- A. Church (1936), "An unsolvable problem of elementary number theory", *Amer. J. Math.* 58, 345–363, Reprinted in *Davis* 1965.
- S. A. Cook (1971), "The complexity of theorem proving procedures", *Proc. 3d ACM STOC*, 151–158.
- M. Davis (1965), *The undecidable. Basic papers on undecidable propositions, unsolvable problems and computable functions*, Raven Press.
- M. Davis (1982), "Why Gödel didn't have Church's Thesis", *Information and Control* 54, 3–24.
- S. Feferman (1977), "Inductive schemata and recursively continuous functionals" In *Logic Colloquium'76*, North-Holland, 373–392.
- S. Feferman (1988), "Turing in the land of  $0(z)$ ". In: *Herken 1988*, 113–147.
- S. Feferman (1990), "Polymorphic typed  $\lambda$ -calculi in a type-free axiomatic framework". In: *Logic and Computation, Contemporary Maths.* 104, AMS, 101–136.
- S. Feferman (1991), "Logics for termination and correctness of functional programs". In *Logic from Computer Science*, MSRI Publications, Springer-Verlag, 95–127.
- J. E. Fenstad (1980), *General recursion theory: An axiomatic approach*, Springer-Verlag.
- J. E. Fenstad / R. Gandy / G. Sacks (1978) (eds.), *Generalized recursion theory II*, North-Holland.
- J. E. Fenstad / P. Hinman (1974) (eds), *Generalized recursion theory*, North-Holland.
- M. Fitting (1987), *Computability theory, semantics and logic programming*, Oxford UP.
- R. M. Friedberg (1957), "Two recursively enumerable sets of incomparable degrees of unsolvability (solution of Post's problem 1944)", *Proc. Nat. Acad. Sci.* 43, 236–238.
- H. Friedman (1971), "Algorithmic procedures, generalized Turing algorithms, and elementary recursion theory". In: *Logic Colloquium'69*, North-Holland, 361–389.
- R. Gandy (1980), "Church's thesis and principles for mechanisms". In: *The Kleene Symposium*, North-Holland, 123–148.
- R. Gandy (1988), "The confluence of ideas in 1936". In: *Herken 1988*, 55–111.
- M. Garey / D. Johnson (1979), *Computers and intractability: A guide to the theory of NP-Completeness*, W. H. Freeman and Co..
- K. Gödel (1986), *Collected Works Volume I : Publications 1929–1936*, Oxford UP.
- K. Gödel (1990), *Collected Works Volume II : Publications 1938–1974*, Oxford UP.
- D. Harel (1987), *Algorithmics: The spirit of computing*, Addison-Wesley.
- R. Herken (1988) (ed.), *The universal Turing machine. A half-century survey*, Oxford UP.
- P. Hinman (1978), *Recursion-theoretic hierarchies*, Springer-Verlag.
- A. Hodges (1983), *Alan Turing: The enigma*, Simon and Schuster.
- A. Kechris / Y. Moschovakis (1977), "Recursion in higher types". In: *Barwise 1977*, 681–737.
- S. C. Kleene (1938), "On notation for ordinal numbers", *J. Symbolic Logic* 3,

- 150–155.
- S. C. Kleene (1952), *Introduction to metamathematics*, North-Holland.
- S. C. Kleene (1959), "Recursive functionals and quantifiers of finite types I", *Trans. A. M. S.* 91, 1–52.
- S. C. Kleene (1959a), "Countable functionals". In: *Constructivity in mathematics*, North-Holland, 81–100.
- S. C. Kleene (1981), "Origins of recursive function theory", *Annals of the History of Computing* 3, 52–67.
- S. C. Kleene / E. Post (1954), "The upper semi-lattice of degrees of recursive unsolvability", *Annals of Math.* 59, 379–407.
- G. Kreisel (1959), "Interpretation of analysis by means of constructive functionals of finite types". In: *Constructivity in mathematics*, North-Holland, 101–128.
- G. Kreisel (1971), "Some reasons for generalizing recursion theory". In: *Logic colloquium'69*, North-Holland, 139–198.
- M. Lerman (1983), *Degrees of unsolvability*, Springer-Verlag.
- W. Maass / T. Slaman (1989), "Some problems and results in the theory of actually computable functions". In: *Logic Colloquium'88*, North-Holland, 79–89.
- J. Mitchell / R. Harper (1988), "The essence of ML", *Proc. 15th ACM / POPL*, 28–46.
- J. Mitchell / G. Plotkin (1984), "Abstract types have existential type", *Proc. 12th ACM / POPL*, 37–51.
- Y. Moschovakis (1984), "Abstract recursion as a foundation for the theory of algorithms". In: *Computation and Proof Theory*, Lecture Notes in Maths. 1104, 289–364.
- P. Odifreddi (1989), *Classical recursion theory*, North-Holland.
- P. Odifreddi (1990) (ed.), *Logic and computer science*, Academic Press.
- R. Platek (1966), *Foundations of recursion theory*, Ph. D. Thesis, Stanford University.
- E. L. Post (1944), "Recursively enumerable sets and their decision problems", *Bull. AMS* 50, 284–316.
- C. Reade (1989), *Elements of functional programming*, Addison-Wesley.
- H. Rogers, Jr. (1967), *Theory of recursive functions and effective computability*, McGraw-Hill.
- G. E. Sacks (1963), *Degrees of unsolvability*, Annals of Math. Studies 55, Princeton.
- G. E. Sacks (1990), *Higher recursion theory*, Springer-Verlag.
- J. Shepherdson (1988), "Mechanisms for computing over arbitrary structures". In: *Herken 1988*, 581–601.
- J. Shepherdson / H. Sturgis (1963), "Computability of recursive functions", *J. for the ACM* 10, 217–255.
- R. Shore (1977), " $\alpha$ -recursion theory". In: *Barwise 1977*, 653–680.
- S. Simpson (1977), "Degrees of unsolvability: a survey of results". In: *Barwise 1977*, 631–652.
- R. Smullyan (1961), *Theory of formal systems*, Annals Math. Studies 47, Princeton.
- R. Soare (1987), *Recursively enumerable sets and degrees*, Springer-Verlag.



- G. Tamburrini (1987), *Reflections on mechanism*, Ph. D. Thesis, Columbia Univ.
- A. S. Troelstra (1977), "Aspects of constructive mathematics". In: *Barwise 1977*, 973–1052.
- J. Tucker / J. Zucker (1988), "Program correctness over abstract data types with error-state semantics", *CWI Monograph* No. 6, Centre for Math and C. S., Amsterdam.
- A. Turing (1936–37), "On computable numbers with an application to the Entscheidungsproblem", *Proc. London Math. Soc.* 42, 230–267; "A correction", *ibid.* 43 (1937), 544–546, Reprinted in *Davis 1965*.
- A. Turing (1939), "Systems of logic based on ordinals", *Proc. London Math. Soc.* 45, 161–228, Reprinted in *Davis 1965*.

## Computers and Mathematics: The Search for a Discipline of Computer Science<sup>1</sup>

MICHAEL S. MAHONEY (Princeton)

In a discussion on the last day of the second NATO Conference on Software Engineering held in Rome in October 1969, Christopher Strachey, Director of the Programming Research Group at Oxford University, lamented that "one of the difficulties about computing science at the moment is that it can't demonstrate any of the things that it has in mind; it can't demonstrate to the software engineering people on a sufficiently large scale that what it is doing is of interest or importance to them".<sup>2</sup> As example he cited the general ignorance or neglect by industry of the recursive methods that computer scientists took to be fundamental to programming. Blaming industry for failing to support research and faulting theorists for neglecting the real problems of practitioners, he went on to explore how the two sides might move closer together.

Strachey's prescription is of less concern here than his diagnosis, which points to an interesting case study in the relation of science to technology in a field thought to be mathematical at heart. His remarks came at a significant point in the history of computing. It marked the end of two decades during which the computer and computing acquired their modern shape. As the title of a recent book on the early computer industry suggests, it was a time of *Creating the Computer*, when the question, "What is a computer, or what should it be?", had no clear-cut answer.<sup>3</sup> By the late 1960's, the main points of that answer had emerged, determined as much in the marketplace as in the

---

<sup>1</sup> Research for this paper was generously supported by the Alfred P. Sloan Foundation through its New Liberal Arts Program.

<sup>2</sup> Peter Naur, Brian Randell, and J. N. Buxton (eds.), *Software Engineering: Concepts and Techniques. Proceedings of the NATO Conferences* (NY: Petrocelli, 1976), 147.

<sup>3</sup> Kenneth Flamm, *Creating the Computer* (Washington, DC: Brookings Institution, 1988).

laboratory. At the same time, a consensus began to form concerning the nature of computer science, at least among those who believed the science should be mathematical.<sup>4</sup> That consensus, reflected in the new category of "computer science" added to *Mathematical Reviews* in 1970, rested in part on theory and in part on experience, if not of the marketplace directly, then of actual machines applied to actual problems.<sup>5</sup>

As the computer was being created, then, so too was the mathematics of computing. When we think of the computer as a machine, it is not surprising that it should be an object of design, where available means and chosen purposes must converge on effective action. We are less accustomed to the idea that a mathematical subject might be a matter of design, that is, the matching of means to ends that themselves are open to choice. Yet, the agenda of the mathematical theory of computation changed as computers and programs grew in size and complexity during the first twenty years. If, as Saunders Mac Lane has said, mathematics is "finding the form to fit the problem", the mathematics of computing began with a search for the problem. Different people made different choices about what was significant, and the mathematics on which they drew varied accordingly. What follows is a first reconnaissance over this shifting terrain.

The electronic digital stored-program computer emerged from the convergence of two separate lines of development each stretching back over several centuries but generally associated with the names of Charles Babbage and

---

4 Not all did. Several recipients of the ACM's Turing Award addressed the question in their award lectures. Although Marvin Minsky ("Form and content in computer science", 1969) agreed that computers are essentially mathematical machines, he decried the trend toward formalization and urged an experimental, programming approach to understanding them. Allen Newell and Herbert Simon ("Computer science as empirical inquiry: Symbols and search", 1975) took an even stronger empirical stand, arguing that computer science is the science of computers and that the limits and possibilities of computing could be determined only through experience in using them. Donald E. Knuth ("Computer programming as an art", 1974) argued that programming was irreducibly a craft skill, which would resist the automation implicit in a mathematization of computer science. See *ACM Turing Award Lectures. The First Twenty Years, 1966-1985* (NY: ACM Press, 1987).

5 The new subject comprised fields taken from various headings. Programming theory, algorithms, symbolic computation, and computational complexity and efficiency had been the province of numerical analysis. From "Information and Communication" came automata theory, linguistics and formal languages, and information retrieval. To these established categories were added adaptive systems, theorem proving, artificial intelligence and pattern recognition, and simulation.

George Boole in the mid-19th century.<sup>6</sup> The first was concerned with mechanical calculation, the second involved mathematics and logic. In coming together, they brought two models of computation: the Boolean algebra of circuits created by Claude Shannon in 1938 and the mathematical logic of Turing machines devised by Alan Turing in 1936. In the early '50s, the new field of automata theory, inspired a decade earlier by the idea of the nervous system as a switching circuit and recently reinforced by the notion of the brain as computer, encompassed the two models at opposite ends of a spectrum ranging from finite deterministic machines to infinite or growing indeterministic machines.<sup>7</sup>

At the one end, beginning with the work of David A. Huffman, E. F. Moore, and G. H. Mealy, switching theory broadened its mathematical scope beyond Boolean algebra by gradually shifting attention from the internal structure of finite-state machines to the patterns of input they can recognize and thus to the notion of a machine as a mapping or partition of semigroups. By 1964, the field of algebraic machine theory was well established, with close links to the emerging fields that were reconstituting universal algebra.<sup>8</sup> At the other end of the spectrum, during the mid-'60s Turing machines of various types became the generally accepted model for measuring the complexity of

---

6 For a fuller sketch and further reading, see M. S. Mahoney, "The History of Computing in the History of Technology", *Annals of the History of Computing* (hereafter *AHC*) 10(1988), 113-25, and "Cybernetics and Information Technology", in: *Companion to the History of Modern Science*, ed. R. C. Olby *et al.* (London-New York: Routledge, Chapman & Hall, 1989), Chap. 34.

7 W. S. McCulloch and W. Pitts, "A logical calculus of the ideas imminent in nervous activity", *Bulletin of Mathematical Biophysics* 5(1943), 115-33. J. von Neumann, "The general and logical theory of automata". In: *Cerebral Mechanisms in Behavior: The Hixon Symposium*, ed. L. A. Jeffries (NY: Wiley, 1951). Robert McNaughton, "The theory of automata, a survey", *Advances in Computing* 2(1961), 379-421.

8 See, for example, J. Hartmanis and R. E. Stearns, *Algebraic Structure of Sequential Machines* (Englewood Cliffs, NJ: Prentice-Hall, 1966), and Paul M. Cohn, *Universal Algebra* (Dordrecht: Reidel, 1965; 2nd rev. ed., 1981). Cf. Alfred North Whitehead, *A Treatise of Universal Algebra* (Cambridge, 1898), I, 29: "[Boole's algebra, characterized by the relation  $a = a + a$ ,] leads to the simplest and most rudimentary type of algebraic symbolism. No symbols representing number or quantity are required in it. The interpretation of such an algebra may be expected therefore to lead to an equally simple and fundamental science. It will be found that the only species of this genus which at present has been developed is the *Algebra of Symbolic Logic*, though there seems no reason why other algebras of this genus should not be developed to receive interpretations in fields of science where strict demonstrative reasoning without relation to number or quantity is required".

computations, a question that shifted attention from decidability to tractability and enabled a classification of problems in terms of the computing resources required for their solution. First broached by Michael O. Rabin in 1959 and '60, the subject emerged as a distinct field with the work of Juris Hartmanis and Richard E. Stearns in 1965 and acquired its full form with the work of Steven Cook and Richard Karp in the early '70s.<sup>9</sup> The field has formed common ground for computer science and operations research, especially in the design and analysis of algorithms.

As the two historical models of computation developed during the 1950s and '60s, they retained their distinctive characteristics. The one stayed close to the physical circuitry of computers, analyzing computation as it went on at the level of the switches. The other stood far away, considering what can be computed in the abstract, irrespective of the particular computer employed. By the mid-50s, however, a new – and, to some extent, unanticipated – complex of questions had arisen in the middle of these extremes. Programming was becoming an activity in its own right, prompting the development of programming languages and compiling techniques to ease the task of writing instructions for specific machines and to make programs transferable from one machine to another. The then dominant application to numerical analysis meant that most such languages would have a mathematical appearance, and the orientation to the programmer meant that they would use symbolic language. But some languages were aimed at insight into computing itself, and they emphasized the manipulation of symbols as opposed to numerical computation.

The explosion of languages over the decade 1955–65, accompanied by the development of general techniques for their implementation and leading to programs of ever greater size and complexity, established all these things as matters of practical fact. In doing so, they challenged computer scientists to give a mathematical account of them. The challenge grew increasingly urgent as problems of cost, reliability, and managerial control multiplied. The call for a discipline of "software engineering" in 1967 meant to some the reduction of programming to a field of applied mathematics.

---

<sup>9</sup> M. O. Rabin, "Speed of computation and classification of recursive sets", *Third Convention of Scientific Societies*, Israel, 1959, 1–2; "Degree of difficulty of computing a function and a partial ordering of recursive sets", *Technical Report No. 1, ONR*, Jerusalem, 1960. J. Hartmanis and R. E. Stearns, "On the computational complexity of algorithms", *Transactions of the AMS* 117(1965), 285–306. S. Cook, "The complexity of theorem proving procedures", *Proc. 3rd ACM Symposium on Theory of Computing*, 1971, 151–58. R. Karp, "Reducibility among combinatorial problems". In: *Complexity of Computer Computations* (NY: Plenum, 1972), 85–104.

At the same time, it was unclear what mathematics was to be applied and where. At first, automata theory seemed to hold great promise. In 1959, building on Stephen Kleene's general characterization of the events that prompt a response from the nerve nets specified by Warren McCulloch and Walter Pitts, Michael Rabin and Dana Scott showed that finite automata defined in the manner of Moore machines accepted the same "regular" sequences of characters (which corresponded to free semigroups) and extended the result to nondeterministic and other finite automata.<sup>10</sup> Such regular expressions constituted the first of Noam Chomsky's hierarchy of phrase structure grammars, the fourth and highest of which corresponded to Turing machines. The suggested link between automata and mathematical linguistics, with its potential application to machine translation, sparked a burst of research. The early '60s saw the creation of pushdown and linear-bounded automata to correspond to the intermediate levels of context-free and context-sensitive grammars. In 1965 formal language theory emerged as a field of computer science independent of the mathematical linguistics from which it had sprung.<sup>11</sup> The new field provided a mathematical basis for lexical analysis and parsing of languages and thus gave theoretical confirmation to techniques such as John Backus' BNF, developed independently for specifying the syntax of Algol.

Even as that work was going on, some writers began to argue that automata theory would not suffice as a mathematical theory of computation.<sup>12</sup> In

---

10 S. C. Kleene, "Representation of events in nerve nets and finite automata". In: *Automata Studies*, ed. J. McCarthy and C. E. Shannon (Annals of Mathematics Studies No. 34, Princeton, 1956), 3-41. M. O. Rabin and D. S. Scott, "Finite automata and their decision problems", *IBM Journal of Research and Development* 3(April 1959), 114-24.

11 Sheila A. Greibach, "Formal languages: Origins and directions", *AHC* 3, 1(1981), 14-41.

12 For example, John McCarthy, in an article to be discussed below, argued that none of the three current (1961) directions of research into the mathematics of computing held much promise of such a science. Numerical analysis was too narrowly focused. The theory of computability set a framework into which any mathematics of computation would have to fit, but it focused on what was unsolvable rather than seeking positive results, and its level of description was too general to capture actual algorithms. Finally, the theory of finite automata, though it operated at the right level of generality, exploded in complexity with the size of current computers. As he explained in another article, "... [T]he fact of finiteness is used to show that the automaton will eventually repeat a state. However, anyone who waits for an IBM 7090 to repeat a state, solely because it is a finite automaton, is in for a very long wait". ("Towards a mathematical science of computation", *Proc. IFIP Congress 62* (Amsterdam: North-Holland, 1963).

principle, the computer was a finite state machine; in practice it was an intractably large finite state machine. Moreover, it was not enough to know that a program is syntactically correct. A program is a function that maps input values to output values and hence is a mathematical object, the structure of which should itself be accessible to expression and analysis. Moreover, programs written in programming languages run on machines and must be translated by means of compilers into the languages of those machines. Both the functional structure of the program and its translation are a matter of semantics, a matter of what the statements of the program, and hence the program itself, mean. Three approaches to semantics emerged in the mid-'60s. The operational approach defined the effective meaning of the language in terms of an abstract machine or definitional interpreter. The deductive approach, introduced by R. W. Floyd in 1967, linked logical statements to the steps of the program, thereby specifying its behavior as well as providing a means of verifying the program.<sup>13</sup> Mathematical semantics aimed at a formal theory that would serve as a means of specification for compilers and as a metalanguage for talking about programs, algorithms and data.

The proposal to make semantics the basis of a mathematical theory of computation came from two sources with different, though complementary emphases. John McCarthy was concerned with the structure of algorithms and how they might be compared with one another. Christopher Strachey spoke about the structure of computer memory and how programs alter its contents. Both men found common ground in a system of functional notation, the lambda calculus, first introduced by Alonzo Church in the early 1930s but subsequently abandoned by him when it did not fulfill his hopes of its serving as a foundation for mathematical logic.<sup>14</sup> The use of the lambda calculus as a metalanguage for programs led to the first construction of a mathematical model for it, and it has subsequently come to be viewed as the "pure" programming language. None of this proceeded smoothly or directly, and it is worth looking at it in a bit more detail.

At the Western Joint Computer Conference in May 1961, McCarthy proposed "A Basis for a Mathematical Theory of Computation".<sup>15</sup> "Computation

---

13 "Attaching Meaning to Programs". In: *Mathematical Aspects of Computer Science* (Proceedings of Symposia in Applied Mathematics, 19; Providence: AMS, 1967), 19–32.

14 J. Barkley Rosser, "Highlights of the history of the lambda calculus", *AHC* 6(1984), 337–49. S. C. Kleene, "Origins of recursive function theory", *AHC* 3(1981), 52–67.

15 Reprinted, with corrections and an added tenth section. In: *Computer Programming and Formal Systems*, ed. P. Braffort and D. Hirschberg

is sure to become one of the most important of the sciences", he began,

This is because it is the science of how machines can be made to carry out intellectual processes. We know that any intellectual process that can be carried out mechanically can be performed by a general purpose digital computer. Moreover, the limitations on what we have been able to make computers do so far clearly come far more from our weakness as programmers than from the intrinsic limitations of the machines. We hope that these limitations can be greatly reduced by developing a mathematical science of computation.<sup>16</sup>

McCarthy made clear what he expected from a suitable theory: first, a universal programming language along the lines of Algol but with richer data descriptions; second, a theory of the equivalence of computational processes, by which equivalence-preserving transformations would allow a choice of among various forms of an algorithm, adapted to particular circumstances; third, a form of symbolic representation of algorithms that could accommodate significant changes in behavior by simple changes in the symbolic expressions; fourth, a formal way of representing computers along with computation; and finally a quantitative theory of computation along the lines of Shannon's measure of information.

McCarthy did not pretend to have met any of these goals, which spanned a broad range of currently separate areas of research. His work on the programming language LISP, however, had suggested a system of formalisms that allowed him to prove the equivalence of computations expressed in them. The formalisms offered means of describing functions computable in terms of base functions, using conditional expressions and recursive definitions. They included computable functionals (functions with functions as arguments), non-computable functions (quantified computable functions), ambiguous functions, and the definition both of new data spaces in terms of base spaces and of functions on those spaces, a feature that Algol, then the most theoretically oriented language, lacked. The system constituted the first part of McCarthy's paper; the second part set out some of its mathematical properties, a method called "recursion induction" for proving equivalence, and a comparison of his system with others in recursive function theory and programming.

In his first presentation of his system in 1960, McCarthy had used a variation of LISP as a metalanguage.<sup>17</sup> He then introduced the lambda calculus, to which he had earlier turned when seeking a notation that allowed the distribu-

---

(Amsterdam, North-Holland, 1963), 33–70.

<sup>16</sup> *Ibid.*, 33.

<sup>17</sup> "Recursive Functions of Symbolic Expressions and Their Computation by Machine", *Communications of the ACM* 3, 4(1960), 184–95.



tion of a function over a list with an indeterminate number of arguments.<sup>18</sup> Concerned primarily with LISP as a working language and satisfied with its metatheoretical qualities, McCarthy did not pursue the further reduction of LISP to the lambda calculus; indeed, he adopted Nathaniel Rochester's concept of *label* as a means of circumventing both the complicated expression and the self-applicative function required by recursive definition in the pure notation. Others, most notably Peter J. Landin, did look to the lambda calculus itself as a metalanguage for programs, seeing several advantages in it.<sup>19</sup> It made the scope of bound variables explicit and thus prevented clashes of scope during substitution (that is one reason why Church designed it). Its rules and procedures for reduction to normal form made it possible to show that two different expressions were equivalent in the sense of having the same result when applied to the same arguments. Moreover, it was type-free, treating variables and functions as equally abstractable entities.

The first property clarified the complexities of evaluating the arguments of a function when their variables have the same name as those of the function.<sup>20</sup> The second property provided analytical insight into the structure of functions, showing how they were constructed from basic functions and allowing transformations among them. In McCarthy's system, it underlay the technique of recursive induction. For example, let integer addition be defined recursively by the conditional equation  $m + n = (n = 0 \rightarrow m, T \rightarrow m' + n^-)$ , where  $m'$  is the successor of  $m$  and  $n^-$  is the predecessor of  $n$ .<sup>21</sup> To show that  $(m + n)' = m' + n$ , let  $g(m, n) = (m + n)' = (n = 0 \rightarrow m, T \rightarrow m' + n^-)' = (n = 0 \rightarrow m', T \rightarrow (m' + n^-)') = (n = 0 \rightarrow m', T \rightarrow g(m', n^-))$ , and  $h(m, n) = m' + n = (n = 0 \rightarrow m', T \rightarrow (m')' + n^-) = (n = 0 \rightarrow m', T \rightarrow h(m', n^-))$ . Whence  $g$  and  $h$  both satisfy the relation  $f = \lambda m. \lambda n. (n = 0 \rightarrow m', T \rightarrow f(m', n^-))$  when substituted for  $f$  and hence are equal.<sup>22</sup>

<sup>18</sup> Interview, 3 December 1990.

<sup>19</sup> P. J. Landin, "The mechanical evaluation of expressions", *Computer Journal* 6(1964), 308–320.

<sup>20</sup> McCarthy and his coworkers had encountered this problem in designing LISP; it came to be called the FUNARG problem. See his account in *History of Programming Languages*, ed. R. Wexelblat (NY: Academic Press, 1981).

<sup>21</sup> The right side of the equation is a conditional expression, which consists of a list of conditional propositions to be evaluated in order from left to right and which takes the value of the consequent of the first proposition of which the antecedent is true. In the above expression, if  $n = 0$ , the value is  $m$ , otherwise ( $T$  is always true) it is  $g(m', n^-)$ ; for example,  $g(3, 2) = g(4, 1) = g(5, 0) = 5$ .

<sup>22</sup> More precisely, in McCarthy's system, they satisfy the relation  $f = \text{label}(f, \lambda m. \lambda n. (n = 0 \rightarrow m', T \rightarrow f(m', n^-)))$ .

The third property opened a link to the particular nature of the stored-program computer and thus fitted McCarthy's rephrasing of his expectations in a second paper, "Towards a Mathematical Science of Computation", delivered at IFIP 62. The entities of computer science consist of "problems, procedures, data spaces, programs representing procedures in particular programming languages, and computers". Once distinguished from problems, defined by the criteria of their solution, the construction of complex procedures from elementary ones could be understood in terms of the well established theory of computable functions. However, there was no comparable theory of the representable data spaces on which those procedures operate. Similarly, while the syntax of programming languages had been formalized, their semantics remained to be studied. Finally, despite the fact that computers are finite automata, "Computer science must study the various ways elements of data spaces are represented in the memory of the computer and how procedures are represented by computer programs. From this point of view, most of the current work on automata theory is beside the point".

McCarthy did not persuade many of his leading American colleagues, who doubted the need for, and feasibility of, a formal semantics, but on this last point he found an ally in Strachey, for whom Landin had been working and who built his contribution to the 1964 Working Conference on Formal Description Languages in Vienna precisely on the question of what goes on in the memory (store) of a computer and on the "essentially computer-oriented" operations of assignment and transfers of control that go on there. In "Toward a Formal Semantics", Strachey worked from the model of a computer's memory as a finite set of  $N$  objects, well ordered in some way by a mapping that assigns to each of them a *name*, or *L-value*. Each object is itself a binary array, which may be viewed as the *value*, or *R-value* associated with the name. A program consists of a sequence of operations applied to names and values to produce values associated with names; in other words, a mapping of names and values into names and values. However the operations are defined abstractly, they reduce to the instruction set of the processor. In principle, one should be able to treat a program as a mathematical object and analyze its structure.

That structure cannot be entirely abstract or syntactical, at least not if it is to meet the most basic requirements of real programming. As an analysis of the assignment command shows, it is necessary to distinguish between the *L-value* and *R-value* of an expression. That is, the command  $\epsilon_1 := \epsilon_2$  requires that the expression on the left be interpreted as a name and that on the right as a value; the two expressions require different evaluations. While one could make that evaluation trivial by restricting the command to allow only primitive names on the left, doing so would sacrifice such features as list-processing in

LISP and Strachey's own CPL. Moreover, the value of a name may extend beyond a binary array to include an area of memory, as in the case, peculiar to the stored-program computer, where it contains executable code. Thus, expressions such as "(if  $x > 0$  then *sin* else *cos*)( $x$ )", meaningless to the mathematical eye, make sense computationally: the variable  $x$  and the procedures for *sin* and *cos* are equally valid values of their names.

To capture the structure of this model of memory, Strachey introduced two operators, "loading" and "updating", which retrieve and store the *R-values* associated with *L-values*. Symbolically, let  $\alpha$  denote an *L-value* and  $\beta$  its associated *R-value*, and let  $\sigma$  denote the "content of the store" or the total set of *R-values* at any given moment. Then "loading", denoted by  $C$ , will be a function of  $\alpha$ , which when applied to  $\sigma$  yields  $\beta$ , that is,  $\beta = (C(\alpha))(\sigma)$ . "Updating", denoted  $U$ , produces a new content  $\sigma'$  through the operation  $(U(\alpha))(\beta', \sigma)$ , where  $\beta'$  is a value compatible with  $\beta$ . Hence, if one treats the "natural" result of an expression  $\epsilon$  as its *L-value*, expressed symbolically as  $\alpha = L(\epsilon, \sigma)$ , then its *R-value* can be obtained by means of the loading operator:  $\beta = R(\epsilon, \sigma) = (C(L(\epsilon, \sigma)))(\sigma)$ . Introduced as functions into the descriptive expressions of a language, Strachey argued, these operators provided a specification of how the results of the expressions should be treated at the level of the computer.

Drawing on Landin's work, Strachey embedded the *L* and *R* functions into the  $\lambda$ -calculus, which he and his collaborators used as a metalanguage for specifying and analyzing the semantics of programming languages.<sup>23</sup> Although they called the enterprise mathematical, it had no underlying mathematical structure to serve as model for the formal system. As Scott insisted when he and Strachey met in Vienna in 1968, their analysis amounted to no more than a translation of the object language into the metalanguage. How the data types and functions of the language were to be constructed mathematically remained an open question. Scott's criticism of Strachey echoed Anil Nerode's reaction to McCarthy's approach.<sup>24</sup> There was no mathematics in it.

<sup>23</sup> P. J. Landin, "The mechanical evaluation of expressions", *Computer Journal* 6(1964), 308–320, develops a "syntactically sugared",  $\lambda$ -less version of Church's notation, which Landin later used to set out a formal specification of the semantics of ALGOL 60. Others undertook to take the approach into the realm of semigroups and categories.

<sup>24</sup> In *Mathematical Reviews* 26(1963), #5766, Nerode wrote that McCarthy had introduced "yet another definition of computability" via conditional expressions and recursive induction. The former is "an arithmetical convenience for handling definition by cases", and the latter, on which McCarthy laid great stress, "is nothing else but the uniqueness of the object defined by a recursive definition". "In the reviewer's opinion", he concluded, "the problem of justifying the title is still open".

Scott had been working in various areas of logic during the '60s, having concluded that none of the areas of theoretical computer science was heading in promising directions.<sup>25</sup> He had gradually formed his own idea of where and how the mathematics entered the picture. He sought the middle ground in the tension inherent in applied mathematics. Mathematics moves in the direction of ever-greater abstraction from the intended application. Yet, the application sets the conditions for the abstraction. The mathematical model must maintain contact with the physical model. The test of practicality always looms over the effort. A mathematical theory of computation addressed to understanding programs has to connect the abstract model to the concrete machine, Scott argued in 1970: "... an *adequate* theory of computation must not only provide the abstractions (what is computable) but also their 'physical' realizations (how to compute them)".<sup>26</sup> The means of realization had been known for some time, he added; what was needed were the abstractions, which could expose the structure of a programming language. "Now it is often suggested that the meaning of the language resides in one particular compiler for it. But that idea is wrong: the 'same' language can have many 'different' compilers. The person who wrote one of these compilers obviously had a (hopefully) clear understanding of the language to guide him, and it is the purpose of mathematical semantics to make this understanding 'visible'. This visibility is to be achieved by abstracting the central ideas into mathematical entities, which can then be 'manipulated' in the familiar mathematical manner".<sup>27</sup>

The mathematical entities derived from the physical structure of the computer. Mathematical semantics concerned *data types* and the *functions* that map them from one to another. The spaces of those functions also form data types. The finite structure of the computer means that some finite *approximation* is needed for functions, which are by nature infinite objects (e.g. mappings of integers to integers). Because computers store programs and data in the same memory, programming languages allowed unrestricted procedures which could have unrestricted procedures as values; in particular a procedure could be applied to itself. "To date", Scott claimed,

no mathematical theory of functions has ever been able to supply conveniently such a free-wheeling notion of function except at the cost of being

---

25 D. S. Scott, "Logic and programming languages" (1976 Turing Award Lecture). In: *ACM Turing Award Lectures*, 47–62.

26 Dana S. Scott, "Outline of a mathematical theory of computation", *Proceedings of the Fourth Annual Princeton Conference on Information Sciences and Systems* (1970); revised and expanded as Technical Monograph PRG-2, Oxford University Computing Laboratory, 1970; 2.

27 *Ibid.*, 3.

inconsistent. The main *mathematical* novelty of the present study is the creation of a proper mathematical theory of functions which accomplishes these aims (consistently!) and which can be used as the basis for the *meta-mathematical* project of providing the "correct" approach to semantics.

One did not need unrestricted procedures to appreciate the problems posed by the self-application facilitated by the design of the computer. Following Strachey, consider the structure of computer memory, representing it mathematically as a mapping of contents to locations. That is, state  $\sigma$  is a function mapping each element  $\ell$  of the set  $L$  of locations to its value  $\sigma(\ell)$  in  $V$ , the set of allowable values. A command effects a change of state; it is a function  $\gamma$  from the set of states  $S$  into  $S$ . Storing a command means that  $\gamma$  can take the form  $\sigma(\ell)$ , and hence  $\sigma(\ell)(\sigma)$  should be well defined. Yet, "[t]his is just an insignificant step away from the self-application problem  $p(p)$  for 'unrestricted' procedures  $p$ , and it is just as hard to justify mathematically".<sup>28</sup>

Recent work on interval arithmetic suggested that one might seek justification through a partial ordering of data types and their functions based on the notion of "approximation" or "informational content". With the addition of an undefined element as "worst approximation" or "containing no information", the data types formed a complete lattice, and monotonic functions of them preserved the lattice. They also preserved the limits of sequences of partially ordered data types and hence were continuous. Scott showed that the least upper bound of the lattice, considered as the limit of sequences, was therefore the least fixed point of the function and was determined by the fixed point operator of the  $\lambda$ -calculus. Hence self-applicative functions of the sort needed for computers had a consistent mathematical model. And so too, by the way, did the  $\lambda$ -calculus for the first time in its history.

Scott's lattice-theoretical model established a rigorous mathematical foundation for the program Strachey had proposed in 1964. Together they wrote "Toward a Mathematical Semantics for Computer Languages", which "covers much the same ground as Strachey ["Toward a Formal Semantics"], but this time the mathematical foundations are secure. It is also intended to act as a bridge between the formal mathematical foundations and their application to programming languages by explaining in some detail the notation and techniques we have found to be most useful".<sup>29</sup> Mathematical semantics formed another sort of bridge as well. It led back to the body of algebraic

---

<sup>28</sup> Ibid., 4–5.

<sup>29</sup> Technical Monograph PRG-6, Oxford University Computing Laboratory, 1971, p. 40; also published in *Proceedings of the Symposium on Computers and Automata*, Microwave Research Institute Symposia Series, Vol. 21, Polytechnic Institute of Brooklyn, 1971.

structures that had provided previous models of computing, but it now spanned the gap between finite-state machines and Turing machines (in the equivalent form of the lambda calculus) by taking account of the random-access, stored program device that embodied them both.

By 1970 computer science had assumed a shape recognized by both the mathematical and the computing communities, and it could point to both applications and mathematical elegance. Yet, it took the form more of a family of loosely related research agendas than of a coherent general theory validated by empirical results. So far, no one mathematical model had proved adequate to the diversity of computing, and the different models were not related in any effective way. What mathematics one used depended on what questions one was asking, and for some questions no mathematics could account in theory for what computing was accomplishing in practice.

In 1969, Christopher Strachey indicated the problem confronting those who looked to computer science for help in addressing the problems of productivity and reliability in the software industry. About a decade later, a committee in the United States reviewing the state of art in theoretical computer science echoed his diagnosis, noting the still limited application of theory to practice.<sup>30</sup> For all the depth of results in computational complexity, "the complexity of most computational tasks we are familiar with – such as sorting, multiplying integers or matrices, or finding shortest paths – is still unknown". Despite the close ties between mathematics and language theory, "by and large, the more mathematical aspects of language theory have not been applied in practice. Their greatest potential service is probably pedagogic, in codifying and given clear economical form to key ideas for handling formal languages". Efforts to bring mathematical rigor to programming quickly reached a level of complexity that made the techniques of verification subject to the very concerns that prompted their development. Mathematical semantics could show "precisely why [a] nasty surprise can arise from a seemingly well-designed programming language", but not how to eliminate the problems from the outset. As a design tool, mathematical semantics was still far from the goal of correcting the anomalies that gave rise to errors in real programming languages.

Another decade later, his successor in the Chair of Computation at Oxford, C. A. R. Hoare, spoke of the mathematics of computing more as aspiration

---

<sup>30</sup> *What Can Be Automated? (COSERS)*, ed. Bruce W. Arden (Cambridge, MA: MIT Press, 1980), 139. The committee consisted of Richard M. Karp (Chair; Berkeley), Zohar Manna (Stanford), Albert R. Meyer (MIT), John C. Reynolds (Syracuse), Robert W. Ritchie (Washington), Jeffrey D. Ullman (Stanford), and Samuel Winograd (IBM Research).

than as reality.<sup>31</sup> He held it as a matter of principle that computers are mathematical machines, computer programs are mathematical expressions, a programming language is a mathematical theory, and programming is a mathematical activity. "These are general philosophical and moral principles, and I hold them to be self-evident – which is just as well, because all the actual evidence is against them. Nothing is really as I have described it, neither computers nor programs nor programming languages nor even programmers".

The sense of anomaly behind such evaluations becomes understandable in light of the historical precedents against which the subject was being viewed. Looking toward a mathematical theory of computation in 1962, McCarthy reached for a familiar touchstone:

In a mathematical science, it is possible to deduce from the basic assumptions, the important properties of the entities treated by the science. Thus, from Newton's law of gravitation and his laws of motion, one can deduce that the planetary orbits obey Kepler's laws.<sup>32</sup>

He extended the analogy at the conclusion of his 1963 article by reference to later successes in mathematical physics:

It is reasonable to hope that the relationship between computation and mathematical logic will be as fruitful in the next century as that between analysis and physics in the last. The development of this relationship demands a concern for both applications and mathematical elegance.<sup>33</sup>

In these historical instances, mathematization had elicited the essential simplicity of an apparently complex world. Newton showed that Kepler's complicated laws followed from the assumption of a simple inverse-square force working according to equally simple laws of motion, and that Galileo's laws of falling bodies could be treated as a limiting case of that model. Hence, pendulums on earth and planets in the heavens move in the same way, and the former can be used to measure the latter both intrinsically (constant of gravity) and extrinsically (marker of time). In an important sense, nineteenth-century mathematical physics merely extended the Newtonian model to other realms of the physical world, even when, as in the case of thermodynamics and electromagnetic theory, the basic laws were substantially different. Those theories tied complicated and diverse phenomena together, drawing them as consequences from a simple mathematical structure. In each case, complexity

---

<sup>31</sup> C. A. R. Hoare, "The Mathematics of Programming", in his *Essays in Computing Science* (Hemel Hempstead: Prentice Hall International, 1989), 352.

<sup>32</sup> "Towards a mathematical science of computation", 21.

<sup>33</sup> "A basis for a mathematical theory of computation", 69.

proved to be accidental, accounted for mathematically by perturbations on a basic solution one moved from the ideal to the real by a form of analytic continuation.

Both theory and experience suggest that, by contrast, complexity is an essential property of computation, to be addressed directly rather than by degrees. Simple structures have provided understanding of simple processes, but they have not readily compounded in any analytic way to give an account of arbitrarily complicated processes.<sup>34</sup> The search for a mathematical structure of computing may well involve a new historical and philosophical structure of mathematization.

---

<sup>34</sup> For this reason, object-oriented programming, however appealing its notion of building complex structures from simple components, lacks any theoretical justification and runs against the grain of experience.





# Global Dimensions of Knowledge: Information, Implementation, and Intertheoretic Relations



## Theories and the Flow of Information

JESUS MOSTERIN (Barcelona)

In order to survive we need to detect, process and store information. Fortunately, the universe is full of it. We feed on this generous flow of cosmic information. Life in general, and our civilization in particular, thrive on this richness of information. Still, our natural capacities for detecting and processing information are limited. But we can extend them artificially. Observation instruments help us to detect more signals. And material and formal instruments like computers and theories help us to better process the available information.

Only coded information is useful information. But information can be encoded in many different ways, some of them more efficient and economical than others. Theories, in particular, are powerful devices for the efficient encoding of information. And this is one of the fundamental roles they play in the fabric of human life.

### *1. How Much Information Is Out There to Flow?*

Immediately after the Big Bang the universe was a very hot and dense gas, nearly homogeneous and in thermal equilibrium. Later on, it fell out of equilibrium. The hot gas expanded and condensed into galaxies, stars and other well structured cosmic systems. The order, structure and thermodynamic information of the universe increased dramatically. Observers and things to be observed became possible. All this would contradict the second law of thermodynamics, if it was not for the presence of the great disequilibrating, namely, the uniform expansion of the universe (of spacetime itself).

In the universe, as in any other system subject to irreversible changes, the entropy has been increasing all the time. Disorder has been increasing all the time. But order has also been increasing. Thermodynamic information is being created all the time. This would be contradictory if we defined thermodynamic order or information as negentropy, i.e., as the negative value of entropy, as Wiener [1961, p. 11] and Brillouin [1962, p. 116, 156] did. Obviously  $S$  and

– $S$  cannot both increase at the same time. But there is no problem if we define thermodynamic order or information as the gap between the actual entropy and the maximum possible entropy:  $Or = S_{\max} - S$ . As long as the maximum possible (or potential – in Leyzar's terminology) entropy increases, both actual entropy and order can continue to grow simultaneously.

The uniform expansion of spacetime is the ultimate source of the disequilibrium and free energy required for building information carrying structures and cosmic systems [Frautschi 1988, p. 16; Layzer 1988, p. 31]. The expansion of the universe and the subsequent creation of disequilibrium has been proceeding at a quicker pace than the degrading and entropy-generating processes pointing towards equilibrium. So the maximum possible entropy of the universe has been increasing more rapidly than its actual entropy, which of course has also been increasing at the same time. The net result has been a spectacular growth of structure and order in the universe.

It is amazing how far from equilibrium and how full of information our actual universe is. In order to become aware of it, it is useful to think of those most entropy-filled of all objects, the black holes. Unfortunately information is not conserved. And the most radical annihilation of information takes place in the formation and growth of a black hole. All matter, all radiation and all objects falling into a black hole disappear irretrievably, together with all their properties and particularities. The huge amount of information necessary to describe all those things is lost for ever.

The theory of black holes is extremely simple. Only two parameters (mass and angular momentum) determine uniquely everything about a spherically symmetric black hole. Specifically its entropy  $S$  is proportional to its surface area  $A$ , which is proportional to the square of the mass of the black hole:

$$S = A \cdot (kc^3/4G\hbar) = m^2 \cdot 2\pi(kcG/\hbar)$$

where  $k$  is Boltzmann's constant,  $c$  is the speed of light,  $G$  is Newton's gravitational constant and  $\hbar$  is Planck's constant.

So it is a relatively straightforward task [Bekenstein 1972; Hawking 1975; Börner 1988; Penrose 1989] to calculate the entropy of a black hole which would include all the mass of the universe. If we assume the standard estimation of  $10^{80}$  baryons (protons and neutrons) for the whole universe, and we suppose that all of them collapse together in a black hole (which inherits their combined mass), we get an entropy per baryon of  $10^{43}$  (in natural units, which make  $k = 1$ ). By contrast, the actual entropy of the observable universe seems to be of around  $10^{10}$  units per baryon (of which  $10^9$  belong to the cosmic background radiation).

The surprising result is that the entropy of the hypothetical black hole

universe would be  $10^{33}$  times the entropy of the actual universe. This immense distance from maximum entropy emphasizes how much ordered and full of information our universe is. It is this vast reservoir of objective information which makes viable the existence of information-crunching creatures like us.

The thermodynamic order results in the building of conspicuous structures, which give out lots of differentiated and finely modulated signals in all directions. We detect some of these signals, whose form in-forms our brains. It is through the processing of this primordial information that we are able to build our images and representations and theories of the ultimate sender of those signals, the universe.

## *2. Compression of Coded Information.*

Thermodynamic order is a very raw kind of information. Only a minuscule portion of it ever gets detected, filtered and encoded. And only coded information is useful information. In the following, we are going to restrict our attention to coded information.

Coded information is able to be more or less efficiently coded. An inefficient, or redundant or too long code can be made shorter, more efficient or compact, can be compressed.

Organic nature makes use of this compressibility of information. The measures of the structural complexity of an organism (even of our brain) give much higher values than the structural complexity of the DNA of the organism. The (practical) information for building the organism has been greatly compressed in the DNA codification. On the other hand, we know that DNA encoding is very far from being an optimally efficient encoding. It contains lots of redundancy in the form of multiple repetitions of the same DNA segments.

The text of a poem or of a song can include several repetitions of the same stanza or refrain, in which case it is possible to spare letters in the specification of that text. It suffices to write the repeated stanza only once and to indicate each of its repetitions by a short mention. If any long name is repeated in the text, it is also possible to shorten it after its first occurrence, with a new saving of characters. So it is often possible to completely specify a text of  $n$  characters by means of only  $m$  characters, where  $m < n$ . The shortened text contains the same information as the complete text, but the shortened text codifies it more compactly, in less letters. This shortening process achieves compression of the information. The more regular, repetitious or symmetrical the text is, the more suitable it is for compression.

The more irregular it is, the less amenable to compression it will be. This resistance to compression is called complexity.

Nowadays the compression of information is a pressing concern in the development of the integrated video-audio-computer systems. These systems can become interactive only if everything (images, sound, data) is stored in digital form. The problem is that the storage of sounds and pictures, and, specially, of images in movement – of TV-style pictures – requires too many bits of memory for the standard storage systems, such as CD-ROMs, at least if the usual encoding is kept.

In the monitor of a computer or the screen of a TV set a still image is represented as a frame of small color points called pixels. The amount of pixels depends on the resolution of the screen. A high resolution screen can have – let us say – one million pixels. If the system chooses from a palette of – say – 256 colors, we need 8 bits ( $= 1$  byte) for the choice of the color of each pixel, as  $\log_2(256) = 8$ . If we add another byte for the intensity value, we get 16 million bits ( $= 2$  million bytes) as the space needed to store the information of a still frame.

In order for the human eye to perceive the impression of continuous movement, we need to show something like 25 still frames per second. That means the memory space needed to store one second of video is 50 million bytes, equivalent to between 30 and 150 diskettes (depending on the diskette format) or a good hard disk of 50 or 60 MB. The largest digital information storage capacity available on standard equipment we find in the CD-ROM. But a CD-ROM has a capacity of about 750 MB, and that is enough for only 15 seconds of high-resolution video. So we see how daunting the difficulties are for the developers of integrated digital video systems. The solution is being looked for in the compression of information.

In many still frames there are homogeneous zones, for example in the background of the picture. The pixels of those zones can be economically specified by default, assuming that all pixels not specifically described are – say – of a faint black color. In the case of movement video it is possible to achieve high levels of compression. Very often each frame is almost identical with the previous one, with only very slight modifications. Perhaps the arm is slowly moving, the rest of the landscape remaining the same. Then it is enough to code for those changes, giving as instruction for the rest of the frame the repetition of the previous information. With tricks like these it is possible to save huge amounts of bits and to compress the information contained in a video in a single CD-ROM. In order for that goal to be achieved, special algorithms are needed for the compression and the depression of the images. And in order for these algorithms to be run quickly enough for

the images to appear smoothly on the screen and in concurrence with the other processes of the system, special co-processors have to be installed in the hardware, like the ones currently being developed by Intel at the David Sarnoff Research Center in Princeton, which are the basis of the so-called DVI systems [Luther 1989].

Any text, any amount of data, any image or video, any melody or sound, can be encoded as a binary sequence, i.e., as a sequence of zeros and ones. This is the way compact disks store music and computers store data, texts and pictures. And soon this will be the standard way of storing movies. Any coded information can be transcoded into a binary sequence. And the simplicity or complexity of the previous message will reappear as simplicity or complexity of the corresponding binary sequence.

So the study of complexity can be restricted, without loss of generality, to the study of the complexity of binary sequences. That is precisely the endeavor of the algorithmic theory of information (or of complexity).

### *3. Algorithmic Information Theory.*

Let us consider the binary sequences A and B whose first digits are:

A: 001 ...

B: 00101110100100001101010001011111011010010111010000011 ...

Suppose both sequences are 3 million digits long. Sequence A can be described or generated by means of the simple algorithm: "write 001 one million times". Sequence B does not seem to be describable in a much shorter way than by just copying the actual sequence in its entirety. The first sequence is highly regular. It is simple. The second one is very irregular. It is complex.

One exact measure of the complexity of a binary sequence is the length of the minimum program that generates that sequence. Of course that measure would be useless, if it was relative to a particular computer or to a particular programming language. Fortunately it is possible to arrive to an absolute value (up to an additive constant), independent of any variation in the hardware or the programming language. Any universal Turing machine will do the job.

A Turing machine is an idealized computer, whose program takes the form of a binary sequence on a potentially infinite tape. The workings of a Turing machine are totally specified by a table, with which the machine itself can be identified. This table can also be coded as a binary sequence. Alan Turing [1937] proved that there are universal Turing machines. A universal Turing machine U is a Turing machine that simulates the behavior of any other Turing machine. Let p be a program or input (a binary sequence), let T be a



Turing machine (coded also as a binary sequence), and let  $T(x)$  be the result or output produced by the machine  $T$  after processing input  $x$  (if such processing ever halts). A universal Turing machine  $U$  is a Turing machine, such that for every binary sequence  $T$  that codes a Turing machine and every binary sequence  $x$  which is a possible input of that machine,  $U$  processes such input exactly as the machine  $T$  would do it, i.e., for every  $T$  and every  $x$  (both binary sequences):

$$U(T, x) = T(x)$$

A universal Turing machine, conveniently programmed, is able to compute any computable function and, especially, can generate any computable binary sequence. Let us now fix a certain universal Turing machine  $\mathcal{U}$ . A program for generating the binary sequence  $x$  is a binary sequence  $p$ , such that, when  $\mathcal{U}$  receives  $p$  as input, it produces  $x$  as output, i.e., such that  $\mathcal{U}(p) = x$ . Each of these programs has a certain length (a certain number of digits – zeroes or ones):  $\text{length}(p)$ . The minimum length of such programs is a precise measure of the complexity of the binary sequence  $x$ . If no program generates  $x$ , we say that the complexity of  $x$  is  $\infty$ .

This measure is univocal up to an additive constant. If, instead of having chosen the universal Turing machine  $\mathcal{U}$ , we would choose a different universal Turing machine, let us say  $U_0$ , then the different measures of complexity so chosen would have coincided asymptotically, i.e., the difference of their values would have always been less than a fixed number  $c$  (which depends only on  $U_0$ ), so that for long enough sequences both measures would have practically coincided. (A more precise statement of this fact is called the invariance theorem of complexity theory).

$\mathcal{K}(x)$ , the complexity of a binary sequence  $x$ , is the length of the minimal program  $p$  which generates  $x$ , if there are such programs which generate  $x$ , and is  $\infty$ , if there are no such programs.

$$\begin{aligned} \mathcal{K}(x) &= \mu n \exists p (n = \text{length}(p) \wedge \mathcal{U}(p) = x) & , \text{ if } \exists p \mathcal{U}(p) = x \\ \mathcal{K}(x) &= \infty & , \text{ if } \neg \exists p \mathcal{U}(p) = x \end{aligned}$$

The function  $\mathcal{K}$  is not computable, but has computable approximations.

The oldest precedent of algorithmic complexity theory can be traced to von Mises' attempts to precise the notion of random binary sequence during the period between the wars. The concepts and ideas typical of the theory appear for the first time at the beginning of the 60's in the work of Ray Solomonoff. Finally in 1965 Andrei Kolmogorov published "Three approaches to the quantitative definition of information", where he defined precisely the notion of complexity – now called in his honor Kolmogorov complexity or  $\mathcal{K}$  – as a

measure of the randomness or individual information of sequences, and proved the invariance theorem. Other mathematicians, like G. Chaitin, P. Martin-Löf, L. Levin, P. Gacs y D. Loveland, have also contributed to the development of the theory. (For a good summary of the theory, see Ming Li & Vitanyi, 1990).

Shannon's concept of information is based on the existence of a set of possibilities or alternatives, provided with an a priori probability distribution, and measures our ignorance of which of these possibilities has materialized. Solomonoff, Kolmogorov and Chaitin, on the contrary, are interested in the informational content of an individual object, without any reference to a set of alternatives. They define the informational content of that object in terms of the difficulty it presents to be described or generated.

If an object is very regular, then it is easy to describe, and so it is simple. If, on the contrary, it is very irregular, then it is difficult to describe, and so it is complex. If it is so complex, that the information it contains cannot be compressed, we say it is random. If it is maximally random, it is chaotic. So the chaos is characterized as the maximum of randomness or complexity, and as the opposite of regularity and simplicity.

#### *4. Mathematical Description as an Encoding Process.*

For information to be defined, we need a well-defined framework, with clear-cut alternatives. Mathematical description can provide such a framework. Placed into such a framework, a raw chunk of reality becomes a system and yields information. It is this rigid framework which creates the conditions for coded information to arise in the first place.

The raw information present in reality has been made available as coded or usable information through the simplifying, idealizing and clarifying process of mathematical description and model building.

In order to get theoretical knowledge of the real world, we force it into the mold of the mathematical structures. The real shape of the Earth is ineffable. But we think of it as a sphere, we model it as a sphere, and so we are able to ask for its radius, and to compute its surface and volume. In successive approximations, we can project more complicated geometrical forms into it and get new and more accurate information.

The mathematical world is fictitious, but objective, well defined, with its own truth of the matter, with its clear sets of alternatives, on which to project the real, but fuzzy world of experience. The raw information present in observation has to be filtered, smoothed out and clad in mathematical or theoretical form in order to become coded information, observational report.

### 5. *Axiomatization as Compression.*

A law or formula compresses the information contained in many observations and historical data.

Solomonoff, a disciple of Carnap, shared his teacher's interest in induction, but looked at the subject with a fresh eye. He pointed out that there is a law-like relationship among a set of observations if and only if the series of their descriptions is not random, i.e., if their regularity makes them compressible. He introduced a general theory of inductive reasoning, based on the idea that a scientific law represents a specially efficient way of compressing the information present in many observational reports, which, once coded as binary sequences, are conceived as the initial segments of an infinite binary sequence, generated by the law. This account is relatively unproblematic for low-level generalizations and phenomenological laws. And, given suitable precautions, it can be extended to laws and theories in general.

The axiomatic method is a method for the compression of information. The axiomatization of a theory is the most efficient way to encode the information contained in its theorems. In some cases each theorem can be conceived as a law, summarizing many real or possible observations. When we compress the information contained in the observation reports, we get the laws, and when we compress the laws, we get the axioms. The axioms can then be thought of as laws of laws, as more efficient encoders. The shortest independent axiom system is then random (if it were not, it would not be the shortest one).

We can think of a theory as a way to compactly summarize all the various sentences it is able to prove, as a program which strongly compresses the information contained in the infinite set of its theorems. The diverse consequences of a theory are coded up by the theory's axioms (and the underlying logic). As the axioms are usually finite in number and small in length, and the consequences are infinite in number and of any length, the compression achieved through successful axiomatization is really stupendous.

The first axiomatization was that of Greek geometry by Euclid. Euclid's axioms compressed the information contained in multiple geometrical propositions previously established by other geometers. Soon it was felt that all the rich previous geometrical literature had become obsolete, as whatever information it contained seemed to be already included in Euclid's theory. Ancient texts were written in perishable materials, like papyrus. Those texts could only survive on condition of being continuously copied. After Euclid's *Elements* the previous geometrical books stopped being copied, and all of them have disappeared. Nevertheless no geometrical information seems to have

been lost in the process. Everything which was in them was already codified in Euclid's theory.

#### *6. Complexity and the Limits of Compression.*

Each binary sequence represents a natural number in the base-two numeration system. And we can define the complexity of a natural number as the complexity of its base-two representation. Most of the binary sequences and most of the natural numbers are very complex, and in this sense they can be said to be random or chaotic. Only one sequence from every thousand (of given length) can be compressed by more than 10 digits.

Even if it is easy to prove that a sequence is not random, and so, that its information is compressible (it is enough to present the corresponding compact algorithm which generates it), it is difficult or impossible to prove that a particular sequence is random or incompressible. So, we can prove that most numbers are random, but we are unable to prove that a particular random number is random.

Most numbers are random and, as they become larger, their complexity grows over all finite bounds. But in a theory of a given degree of complexity it is impossible to prove that a number (or its corresponding binary representation) has a complexity much greater than the complexity of the theory itself. In fact, it is possible to associate with any consistent formal theory which includes elementary arithmetic a constant (a natural number)  $c$  – which depends on the theory –, such that no proposition of the form " $\mathcal{K}(x) > c$ " (where  $x$  is a finite binary sequence and  $\mathcal{K}$  is the complexity) can be proved in the theory. In this sense the theory of computational complexity allows us to obtain incompleteness results similar to those of Gödel. [See Chaitin 1987 and Van Lambalgen 1989].

#### *7. Further Compression Through Multiapplicability.*

Further compression of the information contained in a group of different concrete theories can be achieved through the process of theoretical abstraction, on condition that the systems described by those concrete theories have something interesting in common, i.e., that they share some common structure or form, at least when looked upon from a particular point of view. In that case any of the concrete theories of the group, conveniently made precise, can become an abstract theory by dropping its concrete reference, and becoming open to multiple interpretations. This mathematical core or abstract formal theory is a powerful codifier of all the information contained in the previous concrete

theories, which now become interpretations of their new abstract counterpart. The systems they described are now the multiple realizations or applications of the abstract theory.

An abstract theory is an efficient codifier of the information contained in many different concrete theories. This efficiency is made possible by the fact that the same mathematical core can be applied to very different (and initially unintended) situations.

The main thrust of modern mathematics has gone in the direction of abstraction, and has led to the development of abstract theories, realized or incorporated in many different systems or applications. The concrete theories intended to merely summarize certain arithmetical relationships became in time group theory, field theory and, in general, abstract algebra.

The same has happened with physical theories, but here it was an unintended development, which often surprised the theories' own creators.

It is well known that Newton's theory was successfully applied by Newton himself to many different systems: the solar system, the Earth-Moon system, the pendulum, the falling bodies-Earth system, etc. But later on the theory found realizations unexpected by Newton, like its application by Coulomb to electrostatic and magnetostatic forces.

Equally well known is the case of Maxwell's equations, intended to describe Faraday's electromagnetic field. According to Maxwell's theory a wave of changing electric and magnetic fields propagates through space at a fixed speed. When finally Maxwell was able to calculate the value of this velocity, he found to his surprise that it matched the empirically measured speed of light. This unexpected agreement led him to conclude that light was an electromagnetic phenomenon. Later on Herz discovered radio waves, which happened to reflect, refract and obey Maxwell's equations as light does.

Perhaps less well known is the bizarre story of string theory.

In the twenties Heisenberg had developed the theory of the S matrix. This matrix describes a correlation between the initial and the final states of a quantum system, without making any assumption on what goes on in between. For example, it describes a correlation between the masses and momenta of some particles before and after a collision, without inquiring about the collision itself. This matrix allows us to define a function which assigns a certain probability to each pair of states (the probability that the first state, as initial state, leads to the second state, as final state). The theory proposes to analyze the proprieties of this matrix without trying to take into account the underlying mechanism.

The strong force is one of the fundamental forces of nature. Inside its short range, it is the strongest force of all. It is responsible for maintaining the

protons closely packed together in the nucleus of an atom, easily overcoming the electromagnetic force which tries to drive them apart. It is also responsible for the confinement of the quarks inside the hadrons.

At the end of the sixties, some physicists tried to find directly the mathematical form of an S matrix, which determines the probability that two particles of momenta  $p_1$  and  $p_2$ , after strong interaction, result in two particles with momenta  $p_3$  and  $p_4$ , without knowing the local mechanism of the strong interaction. Under certain simplifying assumptions, Gabrielle Veneziano was able to specify as an integral the solution for such cases. Veneziano's original suggestion was essentially just a guess as to what would happen when strongly interacting particles collided. He didn't have any string picture in mind at that time. Later on, in 1970, Yoichiro Nambu found that this model described the quantized motion of small strings. These were interpreted as hypothetical strings uniting the quarks inside the hadrons. Anyway, the theory was soon abandoned, after 'tHooft discovered the renormalization of gauge theories, which gave new impetus to quantum field theories of the Yang Mills type, specially to the Electroweak Theory and (in the case of the strong interaction) to Quantum Chromodynamics.

Nevertheless, around 1980 the discarded string theory of the strong force was resuscitated, but this time not as a theory of the strong interaction, but as a theory of quantum gravity, and later on even as a theory of everything! It all began some years earlier (in 1974), when Joel Scherk and John Schwarz turned the difficulty of the existence of massless particles in Nambu's model by identifying them with photons and gravitons, which are supposed to be massless anyway, thereby changing completely the intended application of the theory. This implied a tremendous reduction of the intended scales. Planck's length (the one adequate for quantum gravity) is smaller than the nuclear scale by a factor of  $10^{20}$  [Schwarz 1988].

The same abstract theory and the same mathematical structure was now being applied to quite different forces and systems than the ones it was originally devised for.

Many concrete real or fictitious systems can share a common structure or form. The information contained in the many different concrete theories or histories about all those systems can be efficiently compressed into a single abstract theory, the theory of the corresponding abstract structure or form.

Of course compression of information is not always possible. And this implies that axiomatization is not always possible. But to push axiomatization and abstraction to the limit is the quest of theoretical science. This quest is not only esthetically rewarding. It is also the way to achieve an ever more efficient and compact encoding of ever larger amounts of information.

## References

- Bekenstein, J.: 1972. "Black holes and entropy", *Phys. Rev.*, D7, 2333–46.
- Börner, G.: 1988. *The Early Universe*, Berlin-New York, Springer-Verlag.
- Brillouin, L.: 1962. *Science and Information Theory* (2nd ed.), New York, Academic Press.
- Chaitin, G.: 1987. *Algorithmic Information Theory*, Cambridge, Cambridge UP.
- Chaitin, G.: 1987. *Information, Randomness & Incompleteness*, Singapore-London, World Scientific Publishing Co.
- Frautschi, S.: 1988. "Entropy in an expanding universe". In: *Entropy, Information and Evolution*, B. Weber (ed.), Cambridge, Mass, The MIT Press.
- Gacs, P.: 1989. "Review of Chaitin's Algorithmic Information Theory", *Journal of Symbolic Logic* 54, 624–627.
- Hawking, S.: 1975. "Particle creation by black holes", *Commun. Math. Phys.*, 43, p. 199–220.
- Kolmogorov, A.: 1965. "Three approaches to the quantitative definition of information", *Problems in Information Transmission* 1, 1–7.
- Kolmogorov, A.: 1968. "Logical basis for information theory and probability theory", *IEEE Transactions on Information Theory*, vol. IT-14, 662–664.
- Layzer, D.: 1988. "Growth of order in the universe". In: *Entropy, Information and Evolution*, B. Weber (ed.), Cambridge, Mass, The MIT Press.
- Luther, A.: 1989. *Digital Video in the PC Environment*, New York, McGraw Hill.
- Ming Li / Vitanyi, P.: 1990. "Kolmogorov complexity and its applications". In Jan van Leeuwen (ed.): *Handbook of Theoretical Computer Science*. Volume A: *Algorithms and Complexity*. Amsterdam, Elsevier.
- Penrose, R.: 1979. "Singularities and time-asymmetry". In: *General Relativity: An Einstein Centenary*, S. Hawking & W. Israel (eds.), Cambridge, Cambridge UP.
- Penrose, R.: 1989. *The Emperor's New Mind. Concerning computers, minds, and the laws of physics*, Oxford, Oxford UP.
- Shannon, C. / Weaver, W.: 1949. *The Mathematical Theory of Communication*, Urbana, Ill., University of Illinois Press.
- Solomonoff, R.: 1964. "A formal theory of inductive inference, Part I and II", *Information and Control* 7, 1–22 and 224–254.
- Schwarz, J.: 1988. "Interview". In: Davies, P. / Brown, J.: *Superstrings. A Theory of Everything?*, Cambridge, Cambridge UP.
- Turing, A.: 1937. "On computable numbers, with an application to the Entscheidungsproblem", *Proc. London Math. Society*, 2nd series, 42, 116–154.
- Van Lambalgen, M.: 1989. "Algorithmic Information Theory", *The Journal of Symbolic Logic* 54, 1389–1400.
- Weber, B. / Depew, D. / Smith, J. (ed.): 1988. *Entropy, Information, and Evolution. New Perspectives on Physical and Biological Evolution*, Cambridge, Mass, The MIT Press.
- Wiener, N.: 1961. *Cybernetics* (2nd edition), Cambridge, Mass, The MIT Press.

# Structuralism and Scientific Discovery

JOSEPH D. SNEED (Colorado)

## *1. Introduction*

*1.1. Purpose.* A procedure for discovering scientific theories is sketched here. The procedure is one of "genetic modification" ([3],[5]) applied to a "population" of structuralist representations of scientific theories ([1],[2]). These two ideas are separable. Genetic modification might be applied to other representations of scientific theories and other search procedures might be applied to structuralist representations. The motivation for combining them is that both structural representations and genetic modification search procedures may, with some plausibility, be viewed as realistic models of actual scientific practice. Neither of these claims will be defended in detail here. The discussion will be restricted to simple relational theories, though the ultimate objective is to extend the method to theories involving more more complex mathematical entities like real numbers.

*1.2. Structuralist Representation.* The representation of scientific theories employed here will be a variant of the structuralist concept of a specialization net of theory elements ([1], Ch. 4). Implicit in this choice of representation is the view that sets of conceptually related laws (specialization nets), in contrast to single laws (theory elements), are the appropriate "units" of conceptual innovation. That is, conceptual innovation in science – here conceived as the discovery of "theoretical concepts" in the structuralist sense ([1], Sec. 2.3) – is a "holistic" enterprise. One can only arrive at theory elements with theoretical concepts by finding them embedded in a more complex structure – a specialization net.

*1.3. PROLOG Implementation.* For the purpose of machine implementation, a syntactic formulation the usual structuralist apparatus will be provided using the programming language PROLOG. The basic idea of this syntax is that sub-categories of models may be characterized by set-theoretic relations among the values of functions whose domain is the objects of the the category. These



functions may be viewed as queries to data bases which are the models. More intuitively, these functions may be viewed as "experiments" on empirical systems – the models. The choice of PROLOG for a syntactic representation of these ideas is motivated largely by the facility with which it may be employed as a query language for relational data bases.

*1.4. Plan.* To pursue these ideas, I will first indicate how the essential ideas of structuralist specialization nets may be represented in PROLOG (Sec. 2). Then I will characterize the "discovery context" for specializations nets (Sec. 3) and finally I will indicate how genetic modification procedures may be applied to this problem (Sec. 4).

## *2. PROLOG Specialization Nets*

### *2.1. The Space of Empirical Theories*

*2.1.1. Theories as Specialization Nets.* The central philosophical thesis underlying this paper is that the structuralist specialization net is the smallest unit of empirical science whose discovery may be regarded as conceptual innovation. No direct defense of this thesis is offered. (But see [10]). However, the success of the research program described here might be taken as indirect support of the thesis. More precisely, one may consider a "space" of specialization nets that are candidates for theories about a certain range of data. Some of these nets account for the data; some do not. Some employ "theoretical concepts" that go beyond the data; others do not. Conceptual innovation is viewed as "discovering" a specialization net with theoretical concepts that accounts for the data. Here, I will sketch how such a "space" of structuralist specialization nets may be represented in PROLOG. PROLOG representations will be provided for: theory elements consisting of non-theoretical structures, theoretical structures and empirical laws; a specialization relation on theory elements; constraints that operate over the entire specialization net; and the empirical content of a theory net – the set of sets of non-theoretical structures compatible with the laws and constraints. One may view this as a standard structuralist specialization net in which all the theory elements have the same constraints. The conception of "content" is that of DIV-10 in [1], p. 179.

*2.1.2. PROLOG Representation.* The general idea of a PROLOG representation is this. A fixed PROLOG vocabulary will be chosen to be used in representing all theories about a given type relational structures. This fixed vocabulary will consist of distinct sub-vocabularies representing respectively: non-theoretical structures, theoretical structures and constraints. The interpretation

of this vocabulary will of course depend on the specific theory. A syntactic concept of specialization net will be defined so that it is broad enough to include interesting cases which might plausibly represent empirical theories with theoretical concepts as well as many uninteresting cases in which nothing like theoretical concepts can be identified.

*2.1.3. Conceptual Innovation.* The discovery of theoretical concepts – conceptual innovation – will be viewed as the discovery of a theory net in which theoretical concepts appear. A suitably general concept of theory net is thus essential to the claim that theoretical concepts can be genuinely discovered by genetic modification processes. Roughly, we cast our conceptual net widely enough to provide the possibility of (syntactic apparatus for) formulating theories with theoretical concepts. But there is no guarantee that theories formulated with this apparatus will in fact have theoretical concepts. More precisely, the discovery process we contemplate could conceptually innovate – find theories that contain theoretical concepts – but, it is in no intuitive way compelled to do so.

## *2.2. PROLOG Representations of Non-theoretical Structures –Mpp*

*2.2.1. Non-theoretical Structures.* The set of non-theoretical structures  $Mpp[U, r]$  will be represented as a set of PROLOG data bases all sharing a common PROLOG vocabulary.  $M[U, r]$  is the set of all finite relational structures of type integer-vector  $r$  whose individuals are members of the set  $U$ . Intuitively, each member of  $Mpp[U, r]$  is a possible configuration of data.

*2.2.2. Data Representations.* The common PROLOG data vocabulary for  $Mpp[U, r]$  consists of a set of PROLOG atoms  $c_i$  – the names of individuals  $U$  – together with atoms  $dom$  and  $rel_i$  serving as predicates of arity 1 and  $r_i$  respectively, representing the domain and the  $i$ -th relation in members of  $Mpp[U, r]$ . Thus:

D1. (A) A PROLOG data vocabulary for  $Mpp[U, r]$  is  $D = [U, R]$ , where  $U = [c_1, \dots, c_k]$ ,  $R = [ [dom, 1], [rel_1, r_1], \dots, [rel_n, r_n] ]$  and  $c_i$ ,  $dom$  and  $rel_i$  are PROLOG atoms.

(B) If  $D$  is a Prolog data vocabulary for  $Mpp[U, r]$ , then a list of PROLOG facts of the following form is a  $D$ -representation for a single non-theoretical structure (indexed by  $i$ ) in  $Mpp[U, r]$ :

$dom(i, c_{lij}).$

```

.
dom(i,c_nij).

rel_1(i,...r1 places filled with c_i's...).
.
.
.
rel_1(i,...r1 places filled with c_i's...).
.
.
.
rel_n(i,...rn places filled with c_i's...).
.
.
.
rel_n(i,...rn places filled with c_i's...).

```

Thus, a PROLOG data base consisting of a number of parts of the above form represents a set of non-theoretical structures.

### 2.3. PROLOG Representations of Theoretical Structures—Mp

**2.3.1. Theoretical Structures.** In the structuralist account of empirical theories the formal part of "theory element"  $\langle \text{Mpp}, \text{Mp}, \text{M}, \text{C} \rangle$  contains a class of theoretical structures Mp which employ theoretical concepts in addition to those appearing in the non-theoretical structures in Mp. Some minimal formal properties are usually require of these concepts. Roughly, we want to consider here all the possible ways to add a theoretical "super structure" to  $\text{Mpp}[\text{U}, \text{r}]$ . To do this we specify a single, fixed PROLOG vocabulary that will be used to construct all these theoretical super-structures. This fixed vocabulary will be given different "interpretations" by specifying different ways in which its members relate to the data vocabulary and to each other. This vocabulary will subsequently be used to formulate empirical laws – to specify the class of models M in the structuralist theory element. But the content of these laws will always be relative to a specific interpretation of the vocabulary. It is important to understand here that we are considering a "space" consisting of different species of theoretical structures – different Mp's – not different members of the same Mp.

**2.3.2. Auxiliary Vocabulary.** Since an auxiliary vocabulary and a construction based on it is required for both theoretical structures and constraints (C in the structuralist theory element), it is convenient to consider these together. The

specific application to constraints will be considered below. An auxiliary vocabulary is simply a pair  $[G, K]$  of lists of PROLOG atoms cum integers indicating the arity to be given to the atom when it serves as a PROLOG predicate. The list  $G$  contains predicates that will be used for theoretical concepts appearing in laws. The list  $K$  contains predicates that will be used in expressing constraints.

D2. An auxiliary vocabulary is a list:  $A = [G, K]$  such that:

1. The law vocabulary  $G$  is a list of of PROLOG [atom, integers] pairs:  $G = [[p_1, a_1], \dots, [p_K, a_k]]$ . so that  $p_i \neq p_j$ ;
2. For integers  $b, a > 1$  the constraint vocabulary  $K$  is a list  $K = [K_2, \dots, K_b]$  where  $K_i$  is such that:
3. For integer  $a > 1$ , the  $a$ -ary constraint vocabulary  $K_a$  is a list of pairs of PROLOG atoms and integers:  $K_a = [k_1, a+a_1], \dots, [k_m, a+a_m]$ ,. so that  $k_i \neq k_j$  and  $p_i \neq k_j$ .
4. VAR is a finite set of PROLOG variables and 'I', 'I1', 'I2' are PROLOG variables not in VAR

The need to distinguish arity of constraints will become clear below. We will also need a specific finite set of PROLOG variables VAR to be used as "individual" variables and some variables 'I', 'I1', excluded from VAR to be used as "modal" variables.

**2.3.3. Based Procedure Sets.** The concept of a based procedure set will be used to construct procedures used in both laws and constraints. The idea here is that we have two "vocabularies"  $W$  and  $B$ . We assume we have procedures for predicates in  $B$ . They may be extensive, data base predicates (in the case of laws) or previously defined extensive procedures (in the case of constraints). We want to define a set of PROLOG procedures PR for the predicates in  $W$  that are "based on" predicates in  $B$ . Intuitively, the  $B$ -predicates provide the "given data" on the basis of which the  $W$ -predicates are defined. Depending on whether the  $W$ -procedures are "based" on data about single non-theoretical structures or multiple non-theoretical structures, they will 1-ary or  $a$ -ary with  $a > 1$ . 1-ary procedures will be used to construct laws.  $a$ -ary procedures,  $a > 1$  will be used to construct constraints. Sets of procedures PR satisfying suitable conditions we will call  $a$ -ary  $[B, W]$  procedure sets.

**2.3.4.  $a$ -ary  $[B, W]$  Procedure Sets.** The  $B$  vocabulary itself provides the "basic" members of PR. The remainder of PR consists of PROLOG rules whose heads are  $U$  predicates with distinct variables in all their argument places. These rules provide "definitions" for the predicates appearing in their heads. A based procedure set PR will contain exactly one member corresponding to each

member of B and U. Intuitively, PR provides non-trivial definitions for some members of U. Other members of U are, for technical reasons to become apparent later, given trivial definitions. Clearly, in different PR's, the same member of U may be defined in different ways. Intuitively, these are definitions "in terms of" the B predicates. Roughly, this means that only B predicates or U predicates "previously defined" in terms of B predicates may appear in the tails of these definitions. A recursive definition of the set of admissible tails for members of U will make the meaning of 'previously defined' clear. Members of B with distinct variables in all argument places are in PR. The heads (where B predicates count as heads of expressions with null tails) of all members of PR contain distinct free variables in their argument places, though the same predicates may appear in tails of expressions in PR with non-distinct variables in their argument places. Queries are generated from members of PR by instantiation (perhaps null) of variables and constants in these argument places. Each member of a PR may, in this way, provide a number of different queries. These ideas are made precise in the following definition.

D3. If a is a  $>0$  integer and

$B = [[b_{_1},r_1],\dots,[b_{_n},r_n]]$

$W = [[w_{_1},s_1],\dots,[w_{_m},s_m]]$

where  $b_{_i}$  and  $w_{_i}$  are PROLOG atoms and  $r_i, s_i > 0$  integers, then PR is an a-ary  $[B,W]$ -procedure set iff there exists a sets  $TL_0, \dots, TL_m$  so that:

(A) For all  $[b_{_i},r_i]$  in B and for all  $V_{_1}, \dots, V_{_r_i}$  in VAR:

(a)  $b_{_i}(I_1, V_{_1}, \dots, V_{_r_i})$  in  $TL_0$ ;

$b_{_i}(I_a, V_{_1}, \dots, V_{_r_i})$  in  $TL_0$ ;

(b) true in  $TL_0$ ;

(B) For all  $[w_{_j},a_j]$ ,  $1 < k < m$ , in W and for all,  $V_{_1}, \dots, V_{_a_{k-1}}$  in VAR;  
 $w_{_k-1}(I_1, \dots, I_a, V_{_1}, \dots, V_{_a_{k-1}})$  in  $TL_k$ ;

(C) For all  $k < m$ ,

$TL(k-1)$  sub\_set  $TL_k$ ;

(D) For all  $k < m$ , if  $T_1$  and  $T_2$  in  $TL_k$  then:

(a)  $T = \text{not}(T_1)$  in  $TL_k$ ;

(b)  $T = (T_1, T_2)$  in  $TL_k$ ;

(c)  $T = (T_1; T_2)$  in  $TL_k$ .

(E) (a) For all  $[b_{_i},r_i]$  in B and some distinct  $V_{_1}, \dots, V_{_r_i}$  in VAR:

$b_{_i}(I_1, V_{_1}, \dots, V_{_r_i})$  in PR;

$b_{_i}(I_a, V_{_1}, \dots, V_{_r_i})$  in PR;

(b) For all  $[w_{_j},a_j]$  in W there are some distinct  $V_{_1}, \dots, V_{_a_j}$  in VAR and exactly one T in  $TL(j-1)$  so that:

(a) all  $V_{_1}, \dots, V_{_a_j}$  appear in T;

(b)  $w_{_j}(I_1, \dots, I_a, V_{_1}, \dots, V_{_a_j}) \text{ :- } T$  in PR.

$$(F) PR^* = \{ P :- T \text{ in } PR \mid T \neq \text{'true'} \}.$$

**2.3.5. Procedure Sets for Laws.** A procedure set for laws PRG contains PROLOG procedures that can be used to generate queries to the data representation that will appear in laws. Precisely how queries appear in laws will be described immediately below in Sec. 2.4. Since laws all pertain to a single partial potential model we need only 1-ary procedures. The data vocabulary itself provides some of these procedures. Members of R with distinct variables in all argument places are in PRG. These procedures permit the construction of "simple" queries. The law vocabulary provides us with the means of constructing additional queries defined in terms of the the data vocabulary. The remainder of PRG consists of PROLOG rules whose heads are law predicates (members of G) with distinct variables in all their argument places. These rules provide "definitions" for the law predicates appearing in their heads.

D4. If  $D = [U, R]$  is a data vocabulary for  $Mpp[U, r]$  and  $G$  is a law vocabulary then a procedure set for laws (PRG) is a 1-ary  $[G, R]$ -procedure set.

## 2.4. PROLOG Representations of Laws –M

**2.4.1. Procedural Laws.** Our method of representing the empirical laws in theory elements makes use of procedural laws of the form  $p_i < p_j$ . Here ' $p_i$ ' denotes some PROLOG procedure and the notation ' $p_i < p_j$ ' means the set of tuples of individuals for which  $p_i$  succeeds is included in the set for which  $p_j$  succeeds. Such procedural laws characterize classes of models – relational structures represented by PROLOG data bases like those described in Sec. 2.2.2 above. The class of models characterized by  $p_i < p_j$  is just the class of models (represented by PROLOG data bases) for which this set-theoretic relation is true.

**2.4.2. Some Familiar Examples.** Procedural laws are discussed in more detail elsewhere ([12]). Here it must suffice to indicate how they work for some familiar examples. These examples show that some interesting classes of models may be characterized in this way. Just which classes of models may be characterized in this way remains an open question.

$$(E1) \text{ (reflexivity)} \quad \forall x R(x, x).$$

Here let

$$p(X) :- r(X, X).$$

Then the procedural laws  $\text{dom}(X) < p(X)$  and  $p(X) < \text{dom}(X)$  (denoted by  $\text{dom}(X) = p(X)$ ) capture just those models in which  $r$  is reflexive.

(E2) (irreflexivity)  $\forall x \neg R(x,x)$

Here

$p(X) :- \text{not}(r(X,X)).$

works in a law of the form  $\text{dom}(X)=p(X)$  to capture models in which  $r$  is irreflexive.

(E3) (symmetry)  $\forall x \forall y (R(x,y) \rightarrow R(y,x))$

Consider,

$p(X,Y) :- r(X,Y), r(Y,X).$

which succeeds for all and only "symmetric pairs". The procedure law  $r(X,Y) < p(X,Y)$  says intuitively that all pairs in  $r$  are symmetric pairs. It characterizes the same class of models in which  $r$  is symmetric.

(E4) (transitivity)  $\forall x \forall y \forall z (R(x,y) \& R(y,z) \rightarrow R(x,z))$

Consider

$p(X,Y) :- r(X,Z), r(Z,X).$

the transitive closure of  $r$ . The procedure law  $p(X,Y) < r(X,Y)$  simply says that the transitive closure of  $r$  is a sub-set of  $r$  – a defining characteristic of transitive relations.

## 2.5 PROLOG Theory Elements

**2.5.1. Unconstrained Theory Elements.** For purposes of PROLOG representation of specialization nets it is convenient lodge the constraints in the net, rather than in the theory elements. Thus, a theory element becomes a tuple  $\langle \text{Mpp}, \text{Mp}, \text{M} \rangle$ . In the PROLOG representation a theory element is simply a procedure set for laws – PRG – and a set of laws  $L$  sharing the same vocabulary. Nothing else is required. Both the non-theoretical structures  $\text{Mpp}$  and the theoretical structures  $\text{Mp}$  are "buried" in PRG. Thus:

D5. If  $[U, R]$  is a PROLOG data vocabulary for  $\text{Mpp}[U, r]$  and  $G$  is a law vocabulary then a  $[[U, R], G]$ -PROLOG theory element for  $\text{Mpp}[U, r]$  is a list:  $S = [\text{PRG}, L]$  where  $\text{PRG}$  is a 1-ary  $[G, R]$ -procedure set and,  $L$  is a list of procedural laws of the form:  $[q_1 < q_1', \dots, q_i < q_i', \dots, q_m < q_m]$  where  $q_i$  and  $q_i'$  are queries formed by instantiating members of  $G$  with variables from  $\text{VAR}$ .

**2.5.2. Theoretical Concepts.** Note that it is not required that all the procedures defined in  $\text{PRG}$  correspond to queries appearing in some member of  $L$ . Nor is

it required that all queries appearing in  $L$  be derived from procedures with non-trivial definitions in PRG. Interesting, theory elements may satisfy further conditions. For example, some theory elements may be such that all or some of the procedures appearing in  $L$  have non-trivial definitions in PRG. Very roughly, procedures appearing in  $L$  without non-trivial definitions in PRG are to theoretical terms. This rough picture needs the wider context of a specialization net (See below) to be made more precise.

## 2.6. PROLOG Representation of Constraints –C

**2.6.1. Procedure Sets for Constraints.** As in the case of laws (Sec. 2.3.5), a procedure set for constraints PRK contains procedures that may be used to generate queries to the data representation that will appear in constraints. For technical reasons that will become apparent below, it is convenient to group constraints by their arity. An  $a$ -ary constraint is one that imposes conditions on sets of models consisting of exactly  $a$  members, with  $a < \infty$ . In familiar cases  $a$  is a small integer. Thus, we define  $PRK_a$  – a procedure set for  $a$ -ary constraints.

D6. If  $D = [U, R]$  is a data vocabulary for  $Mpp[U, r]$ ,  $G$  is a law vocabulary and  $K_a$  is an  $a$ -ry constraint vocabulary then a procedure set for  $a$ -ary constraints ( $PRK_a$ ) is an  $a$ -ry  $[K_a, R \cup G]$ -procedure set.

Note that the "base vocabulary" for constraints is  $(R \cup G)$  – the data vocabulary together with the law vocabulary. Intuitively, this means that constraint procedures may be defined by PROLOG rules containing both data predicates and "previously defined" law predicates. Exactly how this works will become clear below (Sec. 2.6.2).

**2.6.2. Procedural Constraints.** Analogous to procedural laws, procedural constraints are expressions of the form  $k_i < p_j$  where ' $k_i$ ' is some constraint procedure and ' $p_j$ ' is some law or data procedure. As with procedural laws, the interpretation of  $k_i < p_j$  is that the set of tuples of individuals for which  $k_i$  succeeds is a sub-set of the set of tuples for which  $p_j$  succeeds. The difference in the constraint case is that  $k_i$  will contain "modal indices" – e.g.  $k_i(I1, I2, X, Y)$  – that range over more than one model so that the tuples satisfying  $k_i$  will be determined by properties of all these models while  $p_j$  will contain only one index – e.g.  $p_j(I3, X, Y)$ . Intuitively, the constraint requires that parts of the data base indexed by  $I1$  and  $I2$  be related to that part of the data base indexed by  $I3$ . Whether all interesting constraints can be put in this form is an open question. However, some related work suggest that this is not implausible ([11]).



**2.6.3. Constraint Sets.** We may now define a constraint set for each *a*-ary constraint vocabulary *Ka*.

D7. If  $[U,R]$  is PROLOG data vocabulary for  $Mpp[U,r]$  and *G* is a law vocabulary, then a  $[[U,R],G,Ka]$ -constraint set is a pair:  $[PRKa, Ca]$  where *PRKa* is an *a*-ary procedure set based on  $(G \cup R)$  and *Ca* is a list of *a*-ary procedural constraints the form:  $[k_1 < p_1, \dots, k_i < p_i, \dots, k_n < p_n]$  where *k<sub>i</sub>* is a query formed by instantiating of some member of *Ka* with variables from *VAR*, *p<sub>i</sub>* a query formed by instantiating of some member of  $(R \cup G)$  with variables from *VAR*.

## 2.7. PROLOG Specialization Nets

**2.7.1. Homogenous Theory Nets.** An homogenous PROLOG theory net *N* is simply a set of theory elements together with some constraints. It is "homogenous" in the sense that all theory elements in it employ the same data and auxiliary vocabulary. We shall see shortly that this is sufficient to endow the set of theory elements with a specialization relation so that we have effectively reproduced the structuralist concept of "specialization net".

D8. If  $[U,R]$  is a data vocabulary for  $Mpp[U,R]$  and  $[G,K]$  is an auxiliary vocabulary then *N* is a  $[[U,R],[G,K]]$  homogenous PROLOG theory net iff there exist

$TE = [ [PRG1,L1], \dots, [PRGn,Ln] ]$

$CN = [ CN2, \dots, CNb ]$

$CNi = [ [Ci1, PRKi1], \dots, [Cim, PRKim] ]$

so that:

1.  $N = [ TE, CN ]$

2. all  $[RGPi,Li]$  in *TE* are  $[[U,R],G]$  theory elements;

3. all  $[Cij, PRKij]$  in *CNi* are *i*-ary  $[[U,R],G,Ki]$ -procedural constraints.

Note that there is one set of constraints that operate over the entire net. The structuralist conception of theory net is somewhat more general in allowing that different theory elements may have different constraints related by "specialization". I believe that this generality could be added to the PROLOG representation, but I have not yet considered the details of how to do this.

**2.7.2. Specialization Graphs.** Intuitively, when PROLOG theory element *T'* is obtained from *T*, either by adding non-trivial procedure definitions to *P* or laws to *L*, *T'* is a specialization of *T*. Recall the notation of (D3-F) for non-trivial procedure definitions.

D9. If  $T = [P,L]$  and  $T' = [P',L']$  are  $[[U,R],G]$ -PROLOG theory elements then

$sp(T, T')$  ( $T'$  is a specialization of  $T$ ) iff sub-set( $P^*, P'^*$ ) and sub-set( $L, L'$ ).

If  $N = [E, C]$  is a  $[D, G, K]$ -PROLOG theory net then:

$$S = \{ [T, T'] \mid T, T' \text{ in } E \text{ and } sp(T, T') \}$$

is the specialization graph of  $N$ .

As defined here, PROLOG theory nets come equipped with specialization graphs. One simply looks at the set of theory elements and discovers what specialization relations (if any) exist among them. Clearly, many PROLOG theory nets will have uninteresting or null specialization graphs. PROLOG theory nets of interest will have specialization graphs that satisfy certain further conditions.

*2.7.3. Interesting Specialization Nets.* Consider the following special case.

1)  $S$  has a tree-structure. That is, there is some "basic theory element",  $T_b = [P_b, L_b]$ , at the "top" of the tree so that all  $T$  in  $E$  are specializations of  $T_b$ .

2) There is some proper sub-set  $G_t$  of the law vocabulary  $G$  such that the only members of  $G$  that occur in laws are members of  $G_t$ . Members of  $G_t$  may occur in laws together with members of the data vocabulary  $R$ . There may even be some laws containing only members of  $R$ . Other members of  $G$  may appear in procedures in  $P$  as "aids" in the definition of the procedures of interest in  $G_t$ .

3)  $L_b$  contains only laws containing members of  $G_t$  and  $P_b$  contains only trivial definitions of the members of  $G_t$ .

In this case,  $T_b = [P_b, L_b]$  contains "general laws" that appear in theory elements in  $N$ . Intuitively, the members of  $G_t$  are "theoretical predicates" because there are no non-trivial definitions of them in  $P_b$ . The laws in  $L_b$  are "theoretical laws" because they contain at least one "theoretical query". There are no "non-theoretical laws" containing just data predicates (non-theoretical predicates). This is somewhat like the situation in Newtonian mechanics where the predicates 'mass' and 'force' appear in the basic theory element without any clue as to how their values are to be determined and there are no purely kinematic laws.

*2.7.4. Determination of Theoretical Concepts.* Now consider what happens as we go further down a net of the kind just described. Both the number of non-trivial definitions of theoretical predicates and the number of laws "grow". That is, theoretical predicates, only trivially defined in  $T_b$ , come to be non-trivially defined as we move down the tree. These definitions are always associated in theory elements with law sets including (usually properly so) the laws in  $L_b$ . Intuitively, the definitions correspond to methods of "measuring"

or "determining" the values of the theoretical predicates. The laws – including possibly non-theoretical laws – correspond to the conditions that must obtain in order that these determination methods to yield "valid" results. The growth of the net may occur different ways corresponding to different "branches" of the specialization tree. Generally, the same theoretical predicate may be associated with different definitions and different laws in different branches of the tree. When this happens, intuitively, we have a situation in which the same theoretical predicate may be determined in different ways – depending upon conditions obtained in the data, these conditions being specified by the laws accompanying the procedure definitions in the theory element.

## 2.8. Content of PROLOG Specialization Nets

*2.8.1. Content of Specialization Nets.* Intuitively, the content of a structuralist specialization net is the set of sets of non-theoretical structures that the net countenances as "empirically possible". It is defined in terms of the content of the theory elements that are in it. The content of a structuralist theory element is the set of all sets of non-theoretical structures which can be augmented with a set of theoretical structures so that individual theoretical structures satisfy the laws of the theory element and the entire set satisfies the constraints ([1],DII-15). For present purposes, the most natural criterion for membership in of the content for a specialization net appears to be the following. A set of non-theoretical structures  $D$  is in the content of  $N$  iff there is some way of assigning  $D$  and sub-sets of  $D$  to all theory elements in  $N$  so that: 1) the  $D$  sub-sets assigned to sp-related members of  $N$  are related by set-theoretic inclusion and; 2) assigned sub-sets are in the content of the theory elements to which they are assigned; 3) the sets of theoretical augmentations used to satisfy 2) are sub-set related ([1] DIV-10, p. 179).

*2.8.2. A Content Procedure.* A procedural version of the above requirements on content membership is the following. The content procedure proceeds by examining members of  $D = [n_1, n_2, \dots]$  sequentially. At any point in the procedure, some sub-set of  $D$  (perhaps the null set)  $[n_1, n_2, \dots, n_k]$  will have already been assigned to theory elements in  $N$  in a way that satisfies these requirements. The theoretical augmentations  $[t_1, t_2, \dots, t_k]$  that do this will have been recorded. Intuitively, this amounts to recording values theoretical concepts that have been determined in  $[n_1, n_2, \dots, n_k]$ . The procedure then simply tries to find some way of assigning the next member of  $dk+1$  to the already successfully assigned members. To do this for  $nk+1$ , it first records  $t'_{k+1}$  – whatever (if anything) the constraints of  $N$  together with  $[n_1+t_1, n_1+t_2, \dots, nk+tk]$  entail about the theoretical augmentation of  $nk+1$ . Intuitively, this amounts to "im-

porting" values of theoretical concepts "determined" in  $[n_1, n_2, \dots, n_k]$  into  $n_{k+1}$ . Then it simply searches through  $N$  looking for some theory element  $e$  such that when the law procedures in  $e$  query  $n_{k+1} + t_{k+1}$  they produce results satisfying the laws of  $e$ . These law procedures are then added to the record of the imported values  $t_{k+1}$  to produce  $t_{k+1}$ . Intuitively, it seems clear that  $D$ 's for which this procedure succeeds will satisfy the requirements above. But, at this point, I have no proof of this. It is less clear that this procedure will succeed for every  $D$  that satisfies the above conditions. To make it do this, one might have to allow it to try different orderings of the members of  $D$ . Again I have no proof to offer.

**2.8.3. A PROLOG Procedure for Content.** The above described content procedure may be implemented in PROLOG in the following way – considering only the top-level. Suppose we have a PROLOG data base consisting of a PROLOG data representation for  $Mpp[U, r]$  and a PROLOG data base representation for PROLOG theory net  $N$  of the kind described in Sec. 2.7.3. We want to use the theory net  $N$  to "classify" sub-sets of the partial potential models in the data representation as "inside" or "outside" the content of the theory  $N$ . We may represent sub-sets of the partial potential models in this data base simply as lists of integers – noting that the order that we thereby impose may be significant. With this representation of sub-sets, we may consider the procedure:

```
content(N,[I1]) :- find_model(N,I1).
content(N,[IIL]) :- content(N,L),
constraint(N,[IIL]), find_model(N,I).
```

For  $L$ 's of more than one member, the second "content" rule is called recursively until the right-most member of the list –  $I1$  – is reached. At this point, the first "content" rule applies and "find\_model( $N, I1$ )" is called.

### 3. The Context of Discovery

Here I consider how data about  $Mpp[U, r]$ 's might be presented to a theorizer and what we want of theories produced by a theorizer. Here 'theory' is to be understood as 'specialization net'. First, I suggest that, despite initial counter-intuitive appearances, the appropriate instances of data are sets of  $Mpp[U, r]$ . Motivated by a desire to model actual practice in the natural sciences, I then opt for a concept of data presentation in which only "positive instances" of the theory's content are presented to the theorizer. Next, I consider how one might rank order the performance of theories relative to a given data presentation.

Finally, a theory's "capturing" a data presentation is defined as a kind of optimum with respect to this rank ordering.

### *3.1. Data Presentations*

*3.1.1. Holistic Data Instances.* Our task is to model the process by which empirical scientists, or scientific communities, move from "data" to theories that "fit" the data. How are we to conceive of the "format" in which the data is presented to the theorist? Theories are "about" sets of  $Mpp[U, r]$ 's. Their "content" – what they countenance as empirically possible – is a set of such sets. The most straightforward (but not the only) conception of a data presentation appears to be simply a sequence of sets of  $Mpp[U, r]$ 's. Members of this sequence represent sets of  $Mpp[U, r]$ 's "found in nature that must appear in the content of any theory that "captures" this data.

*3.1.2. Data as Positive Instances.* On this conception of a data presentation, the theorizer "sees" only things that are supposed to be in the content of the theory. It never "sees" things that are not supposed to be there. One could adopt a different view on which the theorizer also "sees" examples of sets partial potential models that "must be" excluded from the content of the theory element. This alternative may be plausible for certain kinds of "theorizing" like the construction of grammars for languages. Linguists may have data both about what counts as an acceptable usage as well as what does not. For experimental sciences like chemistry, one might also claim that we have "data" about what kinds of reactions do not occur. However, the plausibility of the claim for this example is considerably reduced when one tries to describe this "negative data" without resort to some kind of general description – which amounts to offering a kind of theory. For field sciences like astronomy and geology where contrived experiments are not possible, data presentations containing "negative instances" appear clearly implausible. While I have no "knock down" argument against using data presentations containing "negative instances", it does seem reasonable – at least, at first cut – to limit data presentations to those containing only "positive instances".

### *3.2. Preference over Theory Nets*

*3.2.1. A Popperian Criterion.* The kind of presentation just sketched does present a certain technical problem – how to rule out trivially "perfect" theories. For this kind of data presentation, the only kind of mistake a theory can make is to wrongly exclude a member of the presentation from its content. That is, data can only reveal that theories are too exclusive, i.e. too strong. A trivial, maximally inclusive theory would avoid all such mistakes. But, such a

theory is clearly not what a theorizer is after. Data presentations with "negative instances" avoid this problem by being able to reject a maximally inclusive theory because it includes the negative instances. How can we rule out trivially perfect theories? One way is to follow Popper and insist that theories be as strong as possible in the sense of being "maximally exclusive" relative to the data. What we really want is a maximally exclusive theory that avoids all mistakes of wrong exclusion. That is, we want to make the content of the theory as small as possible while, at the same time, avoiding the possibility that "new" data will appear that falls outside the content.

*3.2.2. Minimizing Mistakes.* Of course, it may not be possible to avoid all mistakes of wrong exclusion. We might have to be satisfied with a maximally exclusive theory that simply minimizes wrong exclusions. Whether or not we can ultimately find a theory that avoids all wrong exclusions, it is worth considering how one might order theories in terms of their "failure rate" and "strength". To understand what is at stake here, suppose we have in hand a theory  $N$  that fails on the average  $f\%$  of the time. That is, for any initial part of the data presentation, we will find about  $f\%$  of the mpp's in the presentation outside  $\text{content}(N)$ . Consider first theory  $N'$  which has the same content:

$$\text{content}(N') = \text{content}(N),$$

but a lower failure rate,

$$f' < f.$$

Here  $N'$  reduces the failure rate without changing content. Clearly,

$$N' \text{ better\_than } N.$$

Likewise, when  $N'$  reduces the content without changing the failure rate:

$$\text{content}(N') \text{ proper sub-set } \text{content}(N),$$

$$f' = f,$$

Clearly,

$$N' \text{ better than } N.$$

Naturally, when  $N'$  reduces both the content and the failure rate we still have

$$N' \text{ better than } N.$$

The problematic case occurs when  $N'$  reduces the failure rate, but at the cost of increasing the content:

$$\text{content}(N) \text{ proper sub-set } \text{content}(N'),$$

$$f' < f.$$

Here, we are forced to say which is more important in evaluating theories – "correctness" or "strength". It seems evident that we want to say "correctness" is more important and thus:

$N'$  better than  $N$ .

Intuitively, we are always willing to sacrifice strength for correctness – if we must. But, if another theory element  $N''$  comes along also weaker than  $N$ , but stronger than  $N'$  and with the same failure rate as  $N'$  but stronger than  $N'$ , we will prefer  $N''$  to both  $N$  and  $N'$ .

**3.2.3. PROLOG Data Presentations.** Let us begin to make these ideas more precise. Recall that we have chosen to represent  $Mpp[U, r]$  data with PROLOG facts having an index variable as the first argument indicating a specific member of  $Mpp[U, r]$ . Thus, we may represent sets of partial potential models simply as lists of integers:

$$j = [i_1, \dots, i_m].$$

where the integers range over those appearing in the first place of relations in the data base. This is not entirely satisfactory since the ordering of these lists may not be irrelevant to the success of procedures like "content".

D10. If  $[U, R]$  is a PROLOG data vocabulary for  $Mpp[U, r]$ , and  $D$  is a  $[U, R]$  data representation of length  $Nr$  then

$$p : \{Nr \rightarrow R[\{Nr\}].$$

is a  $[U, R]$  data presentation for  $Mpp[U, r]$ .

Intuitively, each  $p(n) = j_n$  in a data presentation represents an ordered set (list) of  $Mpp[U, r]$  structures that have been "found" in nature. The task of theory discovery is to find the best theory whose content captures this presentation in a sense to be made precise shortly.

**3.2.4. Capturing Data Presentations.** What is desired of a theory relative to a data presentation  $p$ ? Suppose the sequence in our data presentation represents the temporal order of observing its members. What we would like to find at "time"  $n$  is a theory  $N$  that is "best" available for  $p(n)$  and continues to be the best as more data in the presentation is observed. Roughly, we'd like to be able to "quit theorizing and go home" and time  $n$ , confident that we could not find a better theory by working longer. A bit more precisely, our objective at  $n$  is to find the "best" theory available for all initial segments of  $p$ . By 'best' we mean a theory so that any stronger theory has a higher failure rate on some

initial segment of  $p$ . We will say that such theories 'capture  $p$ '.

D11. If  $[U, R]$  is a PROLOG data vocabulary for  $Mpp[U, r]$ ,  $[G, K]$  is an auxiliary vocabulary,  $N$  is  $[[U, R], [G, K]]$  homogenous PROLOG theory net and  $p$  is a  $[U, R]$ -data presentation for  $Mpp[U, r]$   $p$ , then  $N$  captures  $p$  iff for all  $[[U, R], [G, K]]$  homogenous PROLOG theory nets  $N'$  such that

$\text{content}(N')$  sub-set  $\text{content}(N)$ ,

there is some  $n$  in  $\mathbb{N}^+$  so that

$$f(n, N) < f(n, N').$$

Roughly,  $N$  captures  $p$  iff any attempt to strengthen  $N$  will ultimately be "done in" by the data in some initial segment of  $p$ . 'Done in' in the sense that the stronger alternative to  $N$  will produce a higher failure rate. Note that 'capture' here means more than simply 'fits the data'. To capture  $p$ ,  $N$  must be the strongest theory that fits the data  $p(n)$  and continues to fit the data for all  $p(m)$ ,  $m > n$ .

#### 4. Strategies for Discovery

##### 4.1. Discovery by Search

4.1.1. *Complexity Ordering.* One common way of thinking about discovering theories that capture a data presentation is to conceive a theorizer as a procedure for systematic search thorough the "space" of theories [14]. This space may be structured in a number of ways – each providing a different kind of "guidance" for search procedures. The easiest structure to impose on this space is the graph determined by rules for generating theories. Our definition of a PROLOG theory net ((D8), Sec. 2.7.1) can easily be cast into the form of a recursive definition in which more complex theory nets are seen to be constructed from less complex theory nets. This definition would determine, in an obvious way, a complexity ordering and thus a directed graph on the set of theory nets. One might imagine searching this graph to try to find theory nets that capture data presentations.

4.1.2. *Content Ordering.* A little reflection reveals that the complexity graph on theory nets has little immediate relation to our objective. Much more interesting, from our point of view would be the directed graph imposed by the "content" ordering on theory nets [21],[22]. It would be convenient if the content and complexity graphs were related in some simple way – e.g. if

$N$  more complex  $N'$  iff  $\text{content}(N)$  in  $\text{content}(N')$ .

Unfortunately, this appears not to be the case. Just how to provide a purely



syntactical representation of the content ordering of PROLOG theory nets remains an open – and crucial – question.

*4.1.3. Search Strategy.* Supposing we had some syntactical way of representing the content order on theory nets, we might consider a search procedure that worked in the following way. Suppose we are sitting at "time"  $n$  and choose arbitrarily some theory net  $N_n$  with failure rate  $f(n, N_n)$ . We then advance in  $p$  choosing a (possibly) new theory net at each step. We move to  $n+1$  and see how  $N_n$  "does" on  $p(n+1)$ . If  $f(n+1, N_n) \leq f(n, N_n)$  we look for a stronger theory net that will do as well as  $N_n$ . That is we proceed "down" the content ordering to see if we can find a theory net stronger than  $N_n$  with no greater failure rate on  $p(n+1)$ . We stop when we hit a theory net that increases the failure rate and back up to the last  $N'$  so that  $f(n+1, N') \leq f(n+1, N_n)$ . We set  $N' = N_{n+1}$  and move on to  $n+2$ . If  $f(n+1, N_n) > f(n, N_n)$ , we "back-up" the content graph until we find some  $N'$  so that  $f(n+1, N') \leq f(n, N_n)$ . We set this  $N' = N_{n+1}$  and move on to  $n+2$ . The order of exploration of upward and downward paths in the content graph remains unspecified here. I make no claim that the procedure sketched above would, in fact, for all  $p$ , converge to a  $N$  capturing  $p$  – much less that it would do so efficiently. I mention it only to provide an example of what "discovery as search" might look like in this context.

#### *4.2. Discovery by Genetic Modification*

*4.2.1. The Genetic Modification Procedure.* An alternative conception of a theorizer is provided by procedure for "genetic" modification of a population. [11],[6]. On this model, one starts with a "population" of theory nets – a relatively small number of randomly generated members of the theory net space over the same vocabulary:

$$A = [N_1, \dots, N_k].$$

At a given "time"  $n$ , each of these theory nets will be characterized by a failure rate on  $p(n)$ ,  $f(n, N_i)$ . This, together with content comparisons, allows us to rank-order members of  $A$  according to their performance on  $p(n)$ . Discovery proceeds by a process of "genetic modification" of the members of  $A$ . This works in roughly the following way. At periodic intervals in the march through  $p$ , lower ranked theory nets in  $A$  are replaced by new theory nets. These new theory nets are constructed from "parts" of the higher ranked theory nets – sometimes modified in random ways. The modified set is then exposed to the  $p$ -environment again and performance records are generated for this set. Modification occurs again in the light of this performance. The hope is that

the process converges to a set  $A^*$  in which all the (not necessarily distinct) theory nets capture  $p$ . The theory nets in  $A^*$  might turn out to be structurally isomorphic copies of each other. In this case we might be tempted to take a "realist" view of the concepts appearing in them. Effectively, this approach regards "discovery" as a kind of "optimization" problem and attacks the optimization problem with stochastic methods designed to be both more effective and efficient than traditional "hill-climbing" or exhaustive enumeration.

*4.2.2. Application to Scientific Discovery.* The application these methods to scientific discovery has been suggested by Holland et. al [12]. However, in specifics the approach I am suggesting here draws heavily on the work of Reynolds in modeling foraging strategies for pre-ceramic hunters and gatherers in the Valley of Oaxaca, Mexico [8]. In effect, I am suggesting that effective practice in empirical science may be modeled in much the same way as effective practice in the technology of subsistence. The only difference is in the complexity of the entities manipulated. The attractiveness of the genetic model for optimization problems in general is that it provides a way of optimizing rather complex structures without searching the large possibility spaces associated with them. Since theory nets are rather complex structures, this motivation applies here. For modeling the discovery of empirical theories, there is an additional potential attraction. The process of "genetic modification" of theory nets may provide a realistic picture of the intellectual and social processes of scientific discovery. In particular, the role of "special theories" in conceptual innovation – the discovery of new theoretical concepts appears to be illuminated by this approach. In what follows, I would like to pursue this idea by suggesting what a reasonable set of "genetic operators" on theory nets might look like and how they might be interpreted.

*4.2.3. Application to PROLOG Representation.* To understand how genetic modification might work on theory nets, we should think of PROLOG theory nets – as complex list structures. They are simply two member lists – the first member is a list of special theories, the second a list of constraints. For practical reasons, it is convenient to assume these lists to be of the same length in all theory nets – e. g. we always have the same number of theory elements and  $c$  constraints. We may effectively consider shorter lists by countenancing a "trivial" theory element and a "trivial" constraint – those which are satisfied by anything. We may thus think of different theory nets resulting from different ways of "filling-in" the theory element slots and the constraint slots. At the "top-level", genetic modification works by "breaking apart" and "recombining" the lists of theory elements and constraints appearing in the population  $A$ . However, theory elements and constraints are themselves complex list struc-

tures – like theory nets – that may be subjected to the same kind of genetic modification one level lower.

*4.2.4. Genetic Operations on Lists.* Genetic modification of list structures will proceed in roughly the same way at all levels. A small number of simple "elementary genetic operators" will be defined. Some of these will be unary – others will be binary. These operators may be combined sequentially to produce more genetic operators of arbitrary arity. Modifications of theory nets in the population A may produce changes in this population in roughly the following ways. When N in A is modified to produce N', N is replaced by N'. When N1 and N2 are modified to produce N1' and N2', we may replace one or both N1 and N2 by the modified versions. In the case we replace only one "exchange" amounts to "substitution".

*4.2.5. Modeling Scientific Methodologies.* Which theory nets are chosen to be modified and replaced and which genetic operators to apply to the chosen nets will depend on the the failure rate and strength of the theory nets and some "policy" about modification. This choice and modification might be done in a variety of, more and less effective, ways. Any particular way of doing this might naturally be termed a "scientific methodology". Naively, one might suppose that program for genetic discovery must be provided at the outset with a scientific methodology. However this is not so. It is possible to permit the program to "experiment with" different methodologies in a way that it gradually "learns" the most effective one. Reynolds' work on foraging behavior [8], mentioned above, contains this feature.

*4.2.6. Genetic Operations on Whole Theory Nets.* Let us now consider specifically what kind of genetic operators might operate on theory nets. We begin with the top-level modifications of whole theory nets.

PERMUTE ELEMENT. The order of the theory element within a single theory element is permuted.

PERMUTE CONSTRAINT. The order of constraints is is permuted within a single theory net.

These operations appear trivial, but they do serve a technical purpose. For technical reasons, the success of the "content" procedure may depend on the order of the special theories or constraints. These operations effectively assure us that we "have a chance" to discover the correct order.

EXCHANGE ELEMENT. Theory element in two theory nets are exchanged.

EXCHANGE CONSTRAINT. Constraints appearing in two theory nets

are exchanged (leaving the constrained procedures unchanged).

These operations and iterations of them permit all possible reshuffling of the lists of theory nets and constraints. Intuitively, they to discover which theory elements work well together and with which constraints. In particular, one might expect an effective scientific methodology to work with theory element constraint pairs that were related in that the constraints were "on" the same procedures that appeared in the laws.

*4.2.7. Genetic Operations on Theory Elements.* Let us now move one level lower and consider operators at various levels on theory elements within a theory net. The list of theory elements is just a list of pairs [Ps,Ls]. So we may consider an operator that exchanges members, the law lists Ls appearing in these pairs. This has the effect of limited redefinition of the procedures appearing in laws since the law-linked predicates are now paired with different rules defining these predicates. This envisions reshuffling the law lists Ls in toto. However, more subtly would be admitted by reshuffling the law list piecewise.

**EXCHANGE LAW PROCEDURES DEFINITIONS.** Members of law lists Ls appearing in the [Ps,Ls] in a single theory element are exchanged.

Note, this operator allows the kind of migration of laws across theory elements that one expects to see resulting in some "general" laws appearing in all theory in the same theory net. It does this in the case that the Ps's are the same (or nearly so) in both theory elements. In the more general case, it allows for experimentation with different ways of "defining" the concepts used in laws. From this perspective, it is an important aspect of "conceptual innovation". Still more possibilities for conceptual innovation are provided by moving one level further down and modify the procedure definitions appearing in individual Ps's. The means for doing this are technically somewhat complex since they must incorporate the rules of PROLOG syntax in a way that assures that the modified structures are generable from these rules. I will not consider these details here.

**REDEFINE LAW PROCEDURE.** Procedures in a single theory element within a single theory net are redefined by genetic operations on PROLOG rules.

The "redefine law procedure" may introduce genuine novelty into the population of theory nets in that it is not just a "recombination" of pieces already present in structures in the populations A.

**PERMUTE LAWS.** The order of the laws in a single theory element is permuted.

Again, this is an operator lacking intuitive significance, but has a technical purpose with the PROLOG context. Note that the changes in special theories within a single theory element produced by these operators may be propagated to other theory elements in the population A vis the operators described in Sec. 4.2.6.

**4.2.8. Genetic Operations on Constraints.** Finally, let us consider genetic modification on constraints one level lower in the list structure of theory nets. Instead of reshuffling constraints in toto among theory nets, we reshuffle them piecewise. Constraints [GK,K] and [GK',K'] appearing in theory nets N and N' may be simply by exchanging K and K'. As in the case of laws, we may thus redefine, in a limited way, the procedures appearing in the constraint K simply by pairing it with a different GK' appearing in some other theory ne. This envisions moving the entire list of constraints K. As with laws, more subtlety could be introduced by allowing piecewise of the list K.

**EXCHANGE CONSTRAINT PROCEDURE DEFINITION.** Members of constraint lists K and K' appearing in the [GK,K] and [GK',K'] in a different theory nets are exchanged.

As with laws, we may also move one level lower still and modify the constraint procedure definitions appearing in some GK.

**REDEFINE CONSTRAINT PROCEDURE.** Constraint procedures redefined by genetic operations on PROLOG rules.

Here again this introduces genuine novelty into the population.

**4.2.9. Intuitive Considerations.** Intuitively, and very roughly, what goes on with genetic modification of theory nets is this: the presence of an auxiliary vocabulary without fixed "interpretation" allows us to try different ways of introducing essentially new concepts into consideration. New concepts introduced in different situations – different theory elements are identified as being the "same" by the fact that they are "used" in the same way in other processes. In effect, this amounts to turning old-fashioned "operationalism" on its head. Concepts are individuated by their common use rather than by common observation procedures. This "trick" – if it works – goes some way to responding to the reservations Hempel [4] and others have raised about automated discovery of theoretical concepts. It may be viewed as a generalization of the work of Langley, et. al. [7] in the "discovery" of concepts

like collision mechanical mass. The "generalization" simply makes explicit the way in which different ways of determining the same concept must be "combined". Theoretical concepts identified in this way are "considered" by tying them in the context of theories we already have to see if they let us do "better". Genetic operators are just an incremental means of doing this. Usually, conceptual innovations are just incremental modifications of concepts we already have. Radical "conceptual revolution" is not impossible – it may just be very rare. I find this idea intuitively attractive as an account of "scientific progress". Whether it can actually be made to work is another question. This paper may be read as a sketch of a research program to answer this question.

### Bibliography

- [1] Balzer, W. / Moulines, C. U. / Sneed J. D., *An Architectonic for Science: the Structuralist Program*, Dordrecht, Reidel, 1987.
- [2] –, "The Structure of Empirical Science – Local and Global". In: *Logic, Methodology and Philosophy of Science VII*. Eds. R. B. Marcus, G. J. W. Dorn and P. Weingartner, Amsterdam, North-Holland, 1986, 291-306.
- [3] Goldberg, D. E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Reading, MA, Addison-Wesley, 1989.
- [4] Hempel, C. G., "Thoughts on the Limitations of Discovery by Computer". In: *Logic of Discovery and Diagnosis in Medicine* (ed. K. F. Schaffner), Berkeley, CA, University of California Press, 1985, 115-122.
- [5] Holland, J. H., *Adaptation in Natural and Artificial Systems*, Ann Arbor, MI, University of Michigan Press, 1975.
- [6] Holland, J. H. / K. J. Holyoak / R. E. Nesbit / P. R. Thagard, *Induction: Processes of Inference, Learning and Discovery*, Cambridge, MA, The MIT Press, 1987.
- [7] Langley, P. / Simon, H. A. / Bradshaw, G. L. / Zytkow, J. M., *Scientific Discovery: Computational Explorations of the Creative Process*, Cambridge, MA, The MIT Press, 1987.
- [8] Reynolds, R. G., "An Adaptive Computer Model for the Evolution of Plant Collecting and Early Agriculture in the Eastern Valley of Oaxaca". In: *Guila Naqyitz: Archaic Foraging and Early Agriculture in Oaxaca, Mexico*. Ed. K. V. Flannery, New York, NY, Academic Press, 1986.
- [9] Sneed, J. D., *The Logical Structure of Mathematical Physics*. (2nd Ed.), Dordrecht, Reidel, 1979.
- [10] –, "Review: Scientific Discovery by Langley, Simon et. al.", *Science* 236 (1987) 1357-1358.
- [11] –, "Constraints as Intertheoretic Relations", unpublished manuscript, 1983.
- [12] –, "Procedural Syntax for Theory Elements", *Pittsburgh Series in Philosophy of Science* (to appear), 1991.
- [13] Sterling, L. / Shapiro E. Y., *The Art of Prolog: Advanced Programming*

- Techniques*, Cambridge, MA, The MIT Press, 1987.
- [14] Ullman, J. D., *Principles of Database and Knowledge-Base Systems*. vol. 1, Rockville, MR, Computer Science Press, 1988.

## Towards a Typology of Intertheoretical Relations

C. ULISES MOULINES (Berlin)

Intertheoretical relations have been a favourite subject of enquiry in the philosophy of science of the last decades. There is good reason for this. Science is not an amorphous bunch of isolated propositions but rather an organic whole of interrelated theories. Moreover, many epistemological issues considered to be crucial in present-day philosophy of science, like reduction, paradigm-change, and the incommensurability thesis, may be analyzed fruitfully only within the more general framework of intertheoretical relations, and from a particular metatheoretical stance. The present attempt at a typology of intertheoretical relations is based on the structuralist metatheory of empirical science.

Needless to say, it is impossible here to provide even an introductory overview of the elements of the structuralist approach. I can only present here, in a very rough way, some of the essential ideas needed to investigate intertheoretical relations. More details may be found in existing expositions of this approach, in particular in the first chapters of our *Architectonic for Science*<sup>1</sup>.

Structuralism owes its name to its starting point in the reconstruction of science, viz. the methodological proposal to view structures and not statements as the basic units of science. The term "structure" is here understood essentially in the sense of Bourbaki. Scientific theories are conceived as complex structures themselves composed of particular kinds of structures; consequently, intertheoretical relations will be viewed as relations between structures.

In a first step – and we need only this first step to go on to our main subject here – those structures that interest us are models in the sense of formal semantics, i.e. structures satisfying some formulas taken as axioms. A model

---

<sup>1</sup> W. Balzer / C. U. Moulines / J. D. Sneed: *An Architectonic for Science*, Reidel, Dordrecht, 1987.



is a tuple of the form  $\langle D_1, \dots, D_n, R_1, \dots, R_m \rangle$  where the  $D_i$  are "base sets" and the  $R_j$  are constructed out of the  $D_j$  as "echelon-sets" in the sense of Bourbaki. In quantitative science, the  $R_j$  will usually represent metric functions defined on some empirical domains. A theory's identity is provided by a class of models in this sense, i.e. by a class of structures satisfying a given list of axioms. The particular formulation of the axioms chosen is regarded as quite irrelevant, so long as they determine the same class of structures and these are the formal representations of the same domain of applications intended.

Though the particular formulation of the axioms is irrelevant for a theory's identity, the distinction between two general kinds of axioms of a theory is not so. We have to distinguish between the framework conditions or conceptual characterizations on the one hand, and the genuine axioms or fundamental laws on the other. The structures of which only the conceptual characterizations are required, we call "potential models"; they represent, so to speak, the theory's conceptual framework. The structures which in addition satisfy the real laws we call "actual models". We symbolize the first class of structures by " $M_p$ ", the second simply by " $M$ ". Clearly,  $M \subseteq M_p$ .

When I said before that, in a first step, a theory's identity is given by a class of models, this was a somewhat ambiguous description. Actually, one should say that the building block for a theory's identity is a pair  $\langle M_p, M \rangle$  with  $M \subseteq M_p$ . Let us call this modeltheoretic unit a "model-element". According to structuralism there are more components within a theory's identity besides potential and actual models. However, this is all we need for the present discussion. It is also true that everything else we need for a formal analysis of science may be constructed either out of a model-element  $\langle M_p, M \rangle$  or out of relationships between different  $\langle M_p^i, M^i \rangle$ . The latter case is what concerns us now: Different types of intertheoretical relations may be identified according to the different types of formal properties the relationships between several  $\langle M_p^i, M^i \rangle$  appear to have.

Up to now, several kinds of intertheoretical relations have been identified within structuralism: specialization, theoretization, reduction, equivalence, and approximation. Some others with no agreed label may be added. Many case studies of these relations in different disciplines may be found in the literature<sup>2</sup>. However, this wealth of intertheoretical relations has not been investigated from a unifying, systematic point of view. For the kind of comparative classification we envisage, we need a fundamental unit of relationship, a sort

<sup>2</sup> They may be gathered from the list of titles compiled by W. Diederich / A. Ibarra / Th. Mormann, "Bibliography of Structuralism", *Erkenntnis*, 30 (1989).

of "relation atoms". They are what may be called "intertheoretical links" or "links", for short. They are the object of the present investigation.

The general notion of a link is simply that of a relation between several model-elements. At least two different model-elements are set in a relationship and the order does matter, i.e., in general, links are not symmetric relations. The connection between the model-elements given is settled on the purely conceptual level, which, of course, does not imply that their respective laws are excluded from a systematic relationship; but the essential thing is to have the connection between the conceptual frameworks at stake so that, formally, links will be defined on the respective potential models:

*Links:* Let  $E^1 = \langle M_p^1, M^1 \rangle, \dots, E^n = \langle M_p^n, M^n \rangle$  be model-elements. We say that  $\lambda$  is a link between  $E^1, \dots, E^n$  iff:

- (1)  $n > 1$
- (2)  $\exists i, j : M_p^i \neq M_p^j$   
 $1 \leq i, j \leq n$
- (3)  $\emptyset \neq \lambda \subseteq M_p^1 \times \dots \times M_p^n$

In the following,  $\lambda$  will be a variable for links in general. Of course, not all relations satisfying this general definition will be interesting for a metatheoretical analysis. Some of them may be, even for purely formal reasons, completely trivial cases of intertheoretical relations which cannot be expected to have any methodological relevance. An obvious case is a relation which would be extensionally identical with the Cartesian product  $M_p^1 \times \dots \times M_p^n$ . A less obvious case of a trivial link, which is implied by the case just stated, though not conversely, is a link fulfilling the condition

$$M^1 \times \dots \times M^n \subseteq \lambda.$$

A link of this sort, even if it were not completely vacuous in the sense that it may put some restrictions on the potential models, would not rule out any actual model of any of the theories involved. This means that the link in question would not add any further information to the information we already have when stating the fundamental laws of each of the theories; in sum, such a link would be superfluous for these model-elements.

In the following, when we speak of links, we'll assume that they are not superfluous in the sense just explained, that is, the condition

$$\text{Not} : M^1 \times \dots \times M^n \subseteq \lambda$$

is satisfied.

Furthermore, we'll make another simplifying assumption: We'll consider only dyadic links, i.e. links of the form

$$\lambda \subseteq M_p^1 \times M_p^2.$$

They seem to be the most characteristic links in empirical science. Typical intertheoretical relations like reduction, equivalence, or approximation are always relations between only two theories at a time. Moreover, it is usually the case that single links which seem to be  $n$ -adic for  $n > 2$  at a first look, on closer analysis prove to be a combination of several dyadic links. However, it is not clear that this is always the case. In his reconstruction of electrodynamics, Thomas Bartelborth has identified a link between three different theories which is apparently not reducible to a combination of dyadic links<sup>3</sup>. Therefore, we should be cautious here and not rule out the possibility of links of higher complexity, though they appear to be the exception rather than the rule. At any rate, dyadic links are undoubtedly the most significant ones and, furthermore, limiting our consideration to them will simplify the formal aspects of our examination without loss of generality in the argument. In the following, we shall restrict our discussion to non-superfluous, dyadic links.

One claim of this paper is that there are two main types of links in empirical science, which, either individually or in combination, make up all identifiable intertheoretical relations. I cannot prove this claim, and it is also difficult to imagine how it could be proved formally, since it depends on a particular analysis of case studies. At best, it can be made more or less plausible. On the other hand, it should be rather easy to disprove the claim by reconstructing, in an adequate way, some particular example of an inter-theoretic relation which is fundamentally not amenable to a combination of links belonging to one type or the other and which, nevertheless, is relevant for some particular branch of empirical science. Up to now, as I'll indicate in a moment, the analysis of different intertheoretical relations provided by the structuralist program seems to support the claim in question.

The first general type of link may be called "an entailment link", the other one "a determining link". They are quite different in nature: Entailment links are somehow "global", in the sense that their general characterization need not contain any reference to particular concepts of the theories involved – though, of course, they have to appear in the formulation of the statements fixing a particular entailment link in a particular case. On the other hand, determining links are, by definition, those that determine some particular concept of one theory by means of another theory, so that their general characterization form already has to have a place for a particular term. To put it in an intuitive,

---

<sup>3</sup> Cf. Th. Bartelborth: *Eine logische Rekonstruktion der klassischen Elektrodynamik*, Peter Lang, Frankfurt / Main, 1988.

though somewhat misleading way, one could say that entailment links connect laws while determining links connect terms of different theories.

Before we introduce the formal characterizations of these two general categories of links, we need some auxiliary notation to make the symbolization easier. Let  $E^1$  and  $E^2$  be any two model-elements. Then:

- (a)  $E^1 \lambda E^2$  iff  $\lambda$  is a link between  $E^1$  and  $E^2$
- (b)  $x^1 \lambda x^2$  iff  $E^1 \lambda E^2$  and  $x^1 \in M_p^1$  and  $x^2 \in M_p^2$  and  $\langle x^1, x^2 \rangle \in \lambda$ .

Further, take a particular  $x_0 \in M_p$  for a given  $E = \langle M_p, M \rangle$  and let  $t_0$  be either a base set or an echelon-set of  $x_0$ ; we'll write " $t_0 \hat{=} x$ ". We'll say that  $t_0$  is a (primitive) term of  $E$ . Now, take all terms of any  $x \in M_p$  appearing on the same place of the tuple where  $t_0$  appears. We symbolize this class by  $t^a$  and call it "an abstract term" of  $E$ . Clearly,  $t_0 \in t^a$ .

*Entailment links:* Let  $E^1, E^2$  be two model-elements.  $\lambda$  is an entailment link between  $E^1$  and  $E^2$  iff

- (1)  $E^1 \lambda E^2$  and  $M^1 \cap D_I(\lambda) \neq \emptyset$  and  $M^2 \cap D_{II}(\lambda) \neq \emptyset$
- (2) For all  $x^1 \in M_p^1, x^2 \in M_p^2$ : if  $x^1 \lambda x^2$  and  $x^2 \in M^2$  then  $x^1 \in M^1$

In the following, we'll use a special symbol for entailment links: We'll write  $E^2 \eta E^1$  to indicate that we have an entailment link between  $E^1$  and  $E^2$ ; so we'll use  $\eta$  as a variable for entailment links only. Let's briefly discuss the content of the definition of entailment links. The first condition just requires that the link involves some actual models of both theories. The second condition is the essential one: It says that the "stronger" model-element  $E^2$  "implies", in a certain sense, the "weaker"  $E^1$ . This sense of implication is not exactly the same as the usual logical implication between statements. For one thing, no common language is presupposed for  $E^1$  and  $E^2$ . There can be all the meaning variance of the world in the concepts of  $E^1$  and  $E^2$ , respectively, so that there is no way to deduce the axioms of  $E^1$  from those of  $E^2$ , and still we may say that the latter theory entails the former in the structural sense propounded here. Moreover, this notion of entailment is weaker than usual logical implication in an intuitive sense, quite independently of language variance: we don't require for it that whenever the laws of the stronger theory are satisfied those of the weaker one will also be satisfied, but only that this will be the case for those models which are appropriately linked by  $\eta$ . They may be a rather small subclass of  $M^2$  and  $M^1$ , respectively. How many of them will be linked by the entailment link will depend on each particular case. Of course, interesting entailment links in empirical science will be those that, ideally, cover all those actual models of both theories which correspond to intended empirical applications. Now, a methodologically significant feature of entail-

ment links is that they make up the cornerstone, so to speak, of two very important intertheoretical relations: reduction and equivalence. It can be shown formally that all cases of reduction and equivalence of theories reconstructed so far essentially consist of entailment links – one-way links in the case of reduction, two-way links in the case of equivalence<sup>4</sup>. This does not mean that any entailment link between two different theories automatically produces a case of reduction or equivalence, since some further restrictive conditions have to be fulfilled to get an intuitively plausible case of reduction or equivalence. But, at any rate, the entailment link is the essential component.

*Determining links:* Let  $E^1 = \langle M_p^1, M^1 \rangle$ ,  $E^2 = \langle M_p^2, M^2 \rangle$  be two model-elements and let  $t^a$  be a term instantiated in the elements of  $M_p^1$ . We say that  $\lambda$  is a determining link between  $E^1$  and  $E^2$  for  $t^a$  iff

- (1)  $E^1 \lambda E^2$
- (2) For all  $x_1^1, x_2^1 \in M_p^1$ ,  $x^2 \in M_p^2$  and for all  $t_1, t_2 \in t^a$ : if  $t_1 \hat{=}_E x_1^1$  and  $t_2 \hat{=}_E x_2^1$  and  $x_1^1 \lambda x^2$  and  $x_2^1 \lambda x^2$ , then  $t_1 = t_2$ .

In the following let's use  $\delta$  as a variable for determining links only, and let's write  $E^2 \delta [t^a] E^1$  to indicate that  $t^a$  is the particular term of  $E^1$  determined by the link  $\delta$  to  $E^2$ . We may say that the models of  $E^2$  provide a unique determination of the term  $t^a$  of  $E^1$  in the sense that, if two models of  $E^1$  are linked to the same model of  $E^2$ , then the values of  $t^a$  in these two models must be the same. The rest of the parameters in these models may be quite different in value but not so for  $t^a$  if there is a determining link for this term going from  $E^1$  to  $E^2$ . It would be more in accordance with scientific practice to weaken the identity in the consequent of the conditional in (2) into an equivalence relation representing value coincidence up to some given scale or invariance transformation, but for the sake of perspicuity let us keep the simpler version with identity. (The more general case could easily be done with this).

Determining links are very common as the building blocks of intertheoretical relations of all kinds. A very conspicuous sort of determining links, though not the only one, is represented by all so-called "identification links", that is, links which identify the values of a given metric function in one theory with the values of some other function in another theory; for example, identifying mass values in classical collision mechanics with mass values in classical particle mechanics – this notwithstanding the fact that a meaning variance theorist or a historicist philosopher of science may contend, with some degree of plausibility, that the concept of mass in one and the other

---

4 Cf. Balzer / Moulines / Sneed, *op.cit.*, Ch.VI.

theory is different –, or the identification of the mole numbers of thermodynamics with the mole numbers of stoichiometry, and so on. A slightly more complex class of determining links is exemplified by those cases where the value of a given function in one theory is identified with the result of some mathematical operations on the values of several functions in another theory – as when you identify the value of pressure in hydrodynamics with the value of the partial derivative of internal energy with respect to volume, the entropy and the mole numbers being held constant, in thermodynamics. Finally, one should not think that determining links may only concern metric functions. We could plausibly argue, for example, that the non-metric preference relation of decision theory is uniquely determined by some behavioral concepts of a psychological theory.

Now, we may ask whether there is some systematic relationship between entailment and determining links. The proposed definitions of entailment and determining links are stated in such general terms that, from a purely formal point of view, we could always make sure that, given any two model-elements  $E^1$  and  $E^2$ , there are arbitrarily chosen entailment and determining links between them. If no further material criteria are required, we could just construct two arbitrary assignments of models of  $E^1$  into models of  $E^2$  with the properties of an  $\eta$  and a  $\delta$  link. Therefore, in this trivial sense, we could always say that the metatheoretical statement "there is an entailment link between  $E^1$  and  $E^2$ " implies the metatheoretical statement "there is a determining link between  $E^1$  and  $E^2$ ", and conversely.

However, this is certainly not what we mean by a "systematic relationship between entailment and determining links". The interesting question is whether, for a *given* entailment link  $\eta$  between  $E^1$  and  $E^2$ , we may derive, in a natural way, its *corresponding* determining link  $\delta$ , or conversely. That is, we should be able to provide a canonical construction of  $\delta$  out of  $\eta$ , or conversely. Now, if we take only the general definitions propounded here, this is obviously *not* the case, since in the general definition of an entailment link there is no reference whatsoever to particular terms of one or the other theory, and conversely, in the general definition of determining links there is no reference whatsoever to the fundamental laws of both theories. In this sense, therefore, the two categories of links are completely independent, and the logical possibility of having an entailment link without a canonically associated determining link, or viceversa, is certainly given. For example, we could conceive of an entailment relation constituting a reduction relation established in such "global" terms that no determination of the parameters of the reduced theory in terms of those of the reducing theory is assumed. Also, we could imagine a term of one theory being determined by certain terms of

another theory without taking the fundamental laws of either one or the other theory into consideration. This is all right from the point of view of a purely formal analysis and I think it is a good strategy to keep the notions of an entailment and a determining link as separate notions. However, the conclusion changes if we take a rather methodological, and not purely formal stance, based on the analysis of particular examples of intertheoretical relations as well as general pragmatic considerations of plausibility. Then, it appears that entailment and determining links are in some systematic and quite interesting relationships also in those cases (and particularly in those cases) that represent "real-life" examples of intertheoretical relations.

Let's take a concrete and quite simple example: the relation between classical collision mechanics and Newtonian particle mechanics. This example has been studied with much care within the structuralist program and may serve as a paradigmatic case for more complex cases that have also been reconstructed in the literature. In a standard axiomatization of both theories, the primitive metric functions of collision mechanics are velocity and mass while those of Newtonian mechanics are position, mass, and force. The actual models of collision mechanics are characterized by momentum conservation whereas those of Newtonian mechanics are determined by Newton's basic laws.

Now, it can be proved formally that collision mechanics is reducible to Newtonian mechanics in the precise sense of the structuralist explication of reduction<sup>5</sup>. This implies, among other things, that there is an entailment link between both theories; in other words: If a physical system conceived in terms of collision mechanics is linked to a system conceived in terms of Newtonian mechanics and the latter actually fulfills Newton's laws, then the former also satisfies momentum conservation. However, the proof that this is always the case essentially hinges upon the assumption that the pairs of linked models of both theories satisfy the conditions that the mass values in one and the other model are the same and that the value of velocity in the collision-mechanical model equals the value of the derivative of position with respect to time in the Newtonian-mechanical model. Clearly, these are nothing but determining links, that is, if two collision-mechanical models are linked to the same Newtonian model they will coincide at least in their values of mass and velocity. In other words:

" $NPM \eta CCM$ " implies " $NPM \delta[m] CCM$ " and " $NPM \delta[v] CCM$ ".

In principle, this implication does not work the other way round. Nothing

---

5 See Balzer / Moulines / Sneed, *op.cit.*, §VI.3.1.

precludes the possibility that we identify the values of, say, the mass function of a collision-mechanical system with the mass values of a corresponding particle-mechanical system without presupposing that the latter actually satisfies Newton's laws. However, a little reflection based on the pragmatics of science reveals this to be an extremely awkward possibility. If we had good reason to assume, or just suspect, that the particle-mechanical system does *not* fulfill Newton's laws (i.e. that it is only a potential but not an *actual* model of *NPM*), then we would very likely refuse to transfer the mass values of the *NPM*-model to the *CCM*-model. In other words, the transfer of the function determined from one theory to the other will only be accepted if we are entitled to assume that the fundamental laws of the former are satisfied (at least approximately). This means, in this particular example, that we assume the implication

"*NPM*  $\delta[m]$  *CCM*" implies "*NPM*  $\eta$  *CCM*"

also to be valid.

I think that these considerations about the methodological relationship between entailment and determining links are not idiosyncratic of the particular example chosen. The same considerations of pragmatic plausibility can be generalized to any case where entailment and determining links are at stake.

In a sense, therefore, entailment links and determining links are themselves equivalent, at least for those cases of intertheoretical relations that might be taken seriously by scientists. They just correspond to two different, but equivalent, perspectives when conceiving intertheoretical relations – the one might be described as "macroanalytic", the other as "microanalytic".





## *Index of Names*

- Abel, N. H.: 12, 219f  
Abrahão, S. M.: 191  
Aczel, P.: 332  
Adamson, R.: 110  
Adrain, R.: 258f, 275  
Agassi, J.: 167  
Agazzi, E.: 311  
Aho, A.: 338, 345  
Airy, G. B.: 262, 275  
Alagic, S.: 339, 345  
Alas, O. T.: 190  
Albertus Magnus: 194  
Alexander, J. W.: 83  
Ampère, A. M.: 219  
Amsterdamski, S.: 300, 311  
Andersen, K. M.: 211  
Anderson, G.: 46  
Anrich: 313  
Apollonios: 206  
Aquinas: 194  
Aranda, E.: VIII  
Arbib, M. A.: 339, 345  
Archimedes: 206  
Arden, B. W.: 361  
Aristotle: 15, 133, 194  
Arnold, V.: 190  
Arribas, M. J.: VII  
Artin, E.: 12  
Aspray, W.: xif, xvi, 223, 230f  
Atiyah, M. F.: 7, 190  
Axer, A.: 235  
Babbage, Ch.: 350  
Bachem, A.: 109  
Backus, J.: 353  
Bacon, F.: 15, 196  
Baker, T. P.: 344f  
Balcázar, J. L.: 343-345  
Balmer: 233  
Balzer, W.: 35-37, 45, 50, 58, 75f,  
156, 167, 401  
Baron, M.: 85, 211  
Barreau, H.: 223  
Barros, J. A. de: 190f  
Barrow, I.: 211f, 216, 219  
Bartelborth, Th.: 406  
Barwise, J.: 331f, 345  
Bayes, Th.: 198f  
Becker, O.: 280f, 294  
Bekemeier, B.: 281, 294  
Bekenstein, J.: 368, 378  
Bell, J.L.: 190  
Benacerraf, P.: xiiif, 59, 61f, 65, 70,  
76  
Berg, J.: 311  
Berggren, J. L.: 221  
Berkeley, G.: 216f, 230  
Bernard, C.: 206  
Bernoulli, D.: 98f, 108, 198, 200,  
215, 218  
Bernoulli, J.: 196, 198, 213, 216  
Bertrand, J. L. F.: 199, 248, 257,  
260, 270, 275  
Bessel, F. W.: 255, 258f, 264f,  
268, 274f, 281  
Beth, E.: 305, 311  
Bianchi, L.: 180  
Bienaymé, I. J.: 254, 260-262, 264,  
275  
Bigelow, J.: 61, 73, 76, 231  
Birkhoff, G. D.: 12, 205  
Bishop, E.: 9, 221, 223  
Blackley, G. R.: 205, 227  
Bochner, S.: 307, 309, 311  
Bodnar, I. M.: 59  
Bohr, H.: 236  
Bolzano, B.: 218-220, 296, 298,  
304, 311  
Bonevac, D. A.: 54f, 58  
Boole, G.: 101-103, 108, 351  
Boolos, G.: 196, 201

- Boone, W. W.: 320  
 Börner, G.: 368, 378  
 Börsch, A.: 277  
 Borzeszkowski, H.-H.: 300, 311  
 Bos, H. J. M.: 84, 211, 297, 311  
 Bourbaki, N.: 9, 45, 61, 75, 148, 156, 161, 167, 169, 171f, 189f, 403  
 Boutroux, P.: 297, 306, 312  
 Boyer, C. B.: 211, 216  
 Bradshaw, J. L.: 401  
 Braffort, P.: 354  
 Brahe, T.: 197  
 Breger, H.: xv, 81, 86  
 Bridgman, P. W.: 154f  
 Brillouin, L.: 367, 378  
 Briskman, L.: 56, 58  
 Britton: 320  
 Brouwer, L. E. J.: 9, 59, 63, 221, 230  
 Browder, F.: 191  
 Brown, J.: 378  
 Brun, V.: 250  
 Brzezinski, J.: 77  
 Buffon, G. L. L.: 215  
 Bunge, M.: 108  
 Burkhardt, J. C.: 241f, 249  
 Burnside, W.: 17, 24  
 Burrichter, C.: 109  
 Buxton, J. N.: 349  
 Bystrov, P.: 59  
 Cajori, F.: 214  
 Cannon, J. T.: 98, 108  
 Cantor, G.: 9, 17, 73, 210, 220f, 224-226, 228  
 Carboni, A.: 25, 30  
 Carnap, R.: 199, 230, 374  
 Carnot, L.: 217, 297, 308  
 Cartwright, N.: 39, 45, 115, 124-126  
 Carus: 15  
 Casari, E.: 312  
 Cassirer, E.: 32-34, 45, 302, 309, 312  
 Castonguay, Ch.: 312  
 Cauchy, A. L.: 96f, 210, 217-220, 222, 228, 281f, 289, 291-293, 306, 308, 310  
 Cavalieri, B.: 211  
 Cayley, A.: 244  
 Chabert, J.-L.: 254, 257, 261, 275  
 Chaitin, G.: 187, 373, 375, 378  
 Chandra, A. K.: 342, 345  
 Chase, S. U.: 12  
 Chauvenet, W.: 267, 270, 275  
 Chebychev, P.: 243-251, 261  
 Chernac, L.: 239, 241f, 249-251  
 Chihara, C.: 119, 128  
 Child, J. M.: 211  
 Cho, Y. M.: 190  
 Chuaqui, R.: 168, 190  
 Church, A.: 8, 346, 354f, 358  
 Churchland, P. M.: 45  
 Cleave, J. P.: 222  
 Cohen, I. B.: 9, 167, 206f  
 Cohn, P. M.: 351  
 Collins, J.: 214f  
 Condillac, E. B. de: 297  
 Cook, S. A.: 344, 346, 352  
 Corcoran, J.: 104, 106, 109  
 Coriolis, G. G.: 96  
 Corneille, P.: 207  
 Corte, B.: 109  
 Coulomb, Ch. A.: 376  
 Courant, R.: 19  
 Cournot, A. A.: 95, 97, 109  
 Craige, J.: 214f  
 Cramer, H.: 236  
 Crofton, M. W.: 256-258, 265, 268, 275  
 Crowe, M. J.: xf, 205-207, 218  
 Crusius, Ch. A.: 137  
 Currie, G.: 167, 223  
 Czuber, E.: 275  
 d'Alembert, J.: 98  
 Da Costa, N. C. A.: xv, 168, 190f  
 Dalla Chiara, M. L.: 191  
 Dantzig, G. B.: 100, 109  
 Dase, Z.: 244, 249  
 Dauben, J. W.: xi, 210, 221, 223-227  
 Davies, P.: 378  
 Davis, M.: 191, 315, 320, 346  
 Davis, Ph.: 89  
 De Morgan, A.: 103, 109, 257, 278  
 Dedekind, R.: 9, 12, 25, 65, 220,

- 225, 283, 292, 294  
 Dell, J.: 191  
 Dempster, M.: 201  
 Depew, D.: 378  
 Descartes, R.: x, 84, 89, 94f, 109,  
 194, 211, 213, 215, 230, 297  
 Dewey, J.: 15  
 Diamond, H. G.: 232, 235, 248  
 Díaz, J.: 345  
 Dickson, L. E.: 235  
 Diederich, W.: 404  
 Dienger, J.: 264, 275  
 Dieudonné, J.: 41, 45  
 Díez, A.: viif  
 Diophantus: 85, 206, 214  
 Dirac, P. A. M.: 19, 136, 168, 175-  
 182  
 Dirichlet, P. G. L.: 310f  
 Dirksen, E. H.: 293f  
 Dobzhansky, Th.: 117  
 Donkin, W. F.: 256-258, 261, 270,  
 275f  
 Doria, F. A.: xv, 190f  
 Dorn, G. J. W.: 77, 401  
 Dornhoff, L.: 45  
 Dostrovsky, S.: 98, 108  
 Drake, F. R.: 156, 167  
 Dress, A.: 17, 30, 87  
 Dreyfus, H.: 81  
 Dreyfus, S.: 81  
 Dubois, D.: 201  
 Dubuc, E.: 18  
 Dugas, R.: 106, 109  
 Duhem, P.: 35, 46  
 Duillier, F. de: 215  
 Dummett, M.: 66, 76  
 Dunmore, C.: 207  
 Dyck, W. von: 110  
 Dyson, F.: 107, 109  
 Ebbinghaus, H. D.: 156, 159, 161,  
 167  
 Eccles: 68, 77  
 Echeverria, J.: xv  
 Edgeworth, F. Y.: 255f, 260, 262,  
 276  
 Edwards, C. H.: 211, 222  
 Eilenberg, S.: 15, 23, 28, 83, 88  
 Einstein, A.: 6, 136, 141-144, 154,  
 179f  
 Eizagirre, X.: viii  
 Elgin, C. Z.: 33, 46  
 Ellis, R. R.: 255-257, 260, 262f,  
 267, 276  
 Emch, G.: 191  
 Encke, J. F.: 241f, 248, 251, 256,  
 262f, 268-270, 273f, 276  
 Engel, F.: 312  
 Engfer, H. J.: 312  
 Erdős, P.: 236, 248, 250  
 Euclid: 5, 84, 193, 206, 236, 300,  
 303, 374  
 Eudoxus: 135f, 193, 206  
 Euler, L.: 41, 98f, 109, 218, 222,  
 231-233, 235-237, 243, 248,  
 250, 284f, 294, 310f  
 Faltings, G.: 11  
 Farebrother, R. W.: 273, 276  
 Faro, E.: 25, 30  
 Fechner, G. T.: 274  
 Feferman, S.: xi, xvi, 57f, 167, 320,  
 336f, 341f, 346  
 Feingold, M.: 212  
 Fenstad, J. E.: 331, 346  
 Fermat, P. de: 80, 85-88, 211, 213,  
 232,  
 Fernando, T.: 314  
 Ferrero, A.: 273, 276  
 Feyerabend, P. K.: 49, 58  
 Field, H.: 61, 70, 76, 115, 117,  
 119-126, 128, 130f, 151, 167  
 Finetti, B. de: 199f  
 Fitting, M.: 336, 346  
 Flamm, K.: 349  
 Flannery: 401  
 Floyd, R. W.: 354  
 Flum, J.: 156, 159, 161, 167  
 Folkerts, M.: 277  
 Fontenelle, B. de: 205, 207-211,  
 215, 217  
 Forster, E. M.: 107  
 Fourier, J. B. J.: 93, 97-99, 107,  
 109, 263, 292  
 Fowler, D. H.: 221  
 Fraenkel, A.: 9, 147, 156, 159,  
 168, 173, 196, 226  
 Fraïssé, R.: 334

- Francoeur, L. B.: 256  
 Franklin, J.: 70, 76, 133  
 Frautschi, S.: 368, 378  
 Frege, G.: 9, 59, 69, 73, 76, 230, 304  
 French, S.: 191  
 Friedberg, R. M.: 325-327, 333, 346  
 Friedman, H.: 334f, 346  
 Frobenius, G. F.: 7  
 Furtado, A. F.: 191  
 Fuss, P. H.: 243  
 Gabauró, J.: 345  
 Gabriel, G.: 106, 109  
 Gacs, P.: 373, 378  
 Gaidenko, P.: 297, 312  
 Galileo, G.: 70, 135, 139, 141, 154, 197  
 Galois, E.: 12, 29  
 Gandy, R.: 315, 317, 331, 335, 339, 346  
 Garavaso, P.: 123, 131  
 Gardies, L.: 312  
 Garey, M.: 343f, 346  
 Gasking, D. A. T.: 70  
 Gauss, C. F.: 10, 218, 231, 232f, 235, 237, 239-244, 248-264, 267f, 276f, 281, 283, 307  
 Gavroglu, K.: 59  
 Gergonne, J. D.: 95  
 Getschko, D.: 190  
 Gibson, J.: 34  
 Giere, R.: 36, 46  
 Gil, F.: 76  
 Gill, J.: 344f  
 Gillies, D. A.: 76, 207  
 Glaisher, J.: 238, 239, 244, 249, 251  
 Glaisher, J. W. L.: 244, 251, 255f, 260f, 264, 266-268, 271, 273f, 277  
 Gleick, J.: 309, 312  
 Glymour, C.: 50, 52, 58  
 Gödel, K.: 9, 33, 122, 134, 221, 227, 304, 317f, 329, 333, 346  
 Goebbels, J.: 15  
 Goldberg, D. E.: 401  
 Goldblatt, R.: 40, 46  
 Goldman, A. I.: 32f, 38, 46  
 Goldstein, L. J.: 232, 235, 247  
 Goodman, N.: ix, 33, 37, 46, 104, 109  
 Goudaroulis, Y.: 59  
 Grabiner, J. V.: 217-220, 224, 289, 291, 294  
 Grant, H.: 280  
 Grassmann, H.: 14-16, 18, 24f, 30, 296, 298, 300f, 303, 312  
 Grattan-Guinness, I.: xiif, 92f, 96f, 100, 102-104, 109f, 211, 216, 219, 228, 306, 310, 312  
 Greene, H. C.: 206  
 Greibach, S. A.: 353  
 Grothendieck, A.: 7, 12, 16, 18, 23, 29f  
 Grottschel, M.: 109  
 Gudermann, C.: 285, 294  
 Guicciardini, N.: 211, 215  
 Guy, R. K.: 234, 236  
 Haack, S.: 56, 58  
 Hadamard, J.: 81, 235f, 248f  
 Hagen, G. H. L.: 258, 264, 266f, 277  
 Hahn, H.: 312  
 Hahn, L. E.: 46  
 Hall, A. R.: 214f  
 Hall, M. B.: 214  
 Hallet, M.: 31, 46  
 Hamilton, W. R.: 181, 185, 225  
 Hankel, H.: 205, 281f, 288, 294, 311, 312  
 Hanson, W. H.: 196, 201  
 Hardy, G. H.: 116-118, 120, 122f, 125-127, 235f, 249  
 Harel, D.: 339, 342, 345f  
 Hargreave, C. J.: 243f  
 Harper, R.: 341, 347  
 Harrison, D.: 12  
 Hartkämper, A.: 59, 167  
 Hartmanis, J.: 351f  
 Hartshorne, Ch.: 46  
 Hasse, H.: 221  
 Hausdorff, F.: 87f, 175  
 Hawking, S.: 368, 378  
 Heaviside, O.: 15  
 Hegel, G. W. F.: 15, 290, 293f,

- 296, 298f, 312  
 Hein, I.: 90  
 Heisenberg, W.: 376  
 Hellman, G.: 65, 75f, 122, 127-129, 131, 196, 201  
 Helmert, F. R.: 256  
 Hempel, C. G.: 70, 400f  
 Henkin, L.: 150  
 Henrici, O.: 99, 110  
 Hensel: 225f  
 Henstock, R.: 191  
 Herder, J. G.: 280, 294  
 Herken, R.: 346  
 Hersh, R.: IX, XVI, 66, 77, 89  
 Herschel, J.: 257f, 262, 267, 277  
 Hertz, H.: 376  
 Higman: 320  
 Hilbert, D.: 5, 59, 104, 119, 144f, 150, 221, 230, 234  
 Hindenburg, C. F.: 284f, 294  
 Hinman, P.: 330f, 346  
 Hintikka, J.: 70, 77, 190  
 Hipparchus: 136  
 Hirschberg, D.: 354  
 Hirzebruch, F.: 7  
 Hoare, C. A. R.: 361f  
 Hodges, A.: 315, 320, 346  
 Hoffmann, B.: 154  
 Holland, J. H.: 397, 401  
 Holyoak, K. J.: 401  
 Hooker, C. H.: 45  
 Hopcroft, J. E.: 338, 345  
 Horstmann, R. P.: 295  
 Hotelling: 100  
 Howson, C.: x, 31, 46, 200f  
 Hudde, J.: 213  
 Huffman, D. A.: 351  
 Hume, D.: 133, 137, 195, 198, 297  
 Hurewicz, W.: 26f  
 Husserl, E.: 61  
 Huygens, C.: 214f  
 Ibarra, A.: xv, 404  
 Ingham, A. E.: 236  
 Inhetveen, R.: 109  
 Irvine, A. D.: XII, XVI, 62f, 77  
 Israel, W.: 378  
 Ivory, J.: 254, 256, 258, 277  
 Jacobi, C. G. J.: 185, 243  
 Jacobson, N.: 12  
 Jahnke, H.-J.: XI, 283, 294, 312  
 Janelidze: 13  
 Jauch, J. M.: 191  
 Jech, T. J.: 157, 167  
 Jeffries, L. A.: 351  
 Jevons, W. S.: 101-103, 110  
 Johnson, D.: 343f, 346  
 Jordan, C.: 12  
 Joyal, A.: 12  
 Kalmar, L.: 155  
 Kaluza, T.: 181  
 Kambartel, F.: 311  
 Kan, D. M.: 12, 16  
 Kant, I.: 32, 59, 137f, 197, 230, 280-283, 293-300, 302, 312  
 Kaplansky, I.: 12  
 Karasmanis, V.: 193, 201  
 Karp, R. M.: 352, 361  
 Kater, H.: 264  
 Kechris, A.: 331, 336, 346  
 Keisler, H. J.: 221, 225-227  
 Kelley, J. L.: 88  
 Kelly, G. M.: 30  
 Kelvin, Lord: 99, 102, 107, 110  
 Kennison, J. F.: 12  
 Kepler, J.: 141, 154, 197, 233  
 Keynes, J. M.: 199  
 Kim, J.: 70, 77  
 Kitcher, Ph.: xIf, xVI, 53-55, 59, 61f, 68, 74, 76f, 119, 223, 230f, 299, 312  
 Kleene, S. C.: 314f, 317, 323, 325, 327-330, 335f, 346f, 353f  
 Klein, F.: 33, 86f, 89, 110, 168, 174, 178, 181f, 191, 307, 310, 312  
 Knobloch, E.: XI, 107, 110, 267, 277  
 Knorr, W.: 221  
 Knuth, D. E.: 350  
 Kobayashi, S.: 191  
 Kochen, S.: 226  
 Kolchin, E.: 12  
 Kolmogorov, A.: 187, 372f, 378  
 König, G.: 109, 313  
 Kötter, R.: 109  
 Krajewski, W.: 50, 59, 305, 312

- Krantz, D.: 71, 77, 119-123, 191  
 Kreisel, G.: 159, 163, 167, 330, 332-334, 337, 347  
 Kripke, S.: 58, 332  
 Krivine, J. L.: 159, 163, 167, 191  
 Kronecker, L.: 221  
 Kucera, A.: 326  
 Kuhn, Th. S.: 48f, 59, 215  
 Kuratowski, C.: 10  
 Kyburg, H.: 115, 123-125  
 l'Hôpital, G. de: 208f, 215, 218  
 L'Huilier, S. de: 217  
 Lacombe, D.: 334  
 Ladner: 344  
 Lagrange, J. L.: 12, 41, 103, 217-219, 281f, 285f, 288-295, 308, 310  
 Lakatos, I.: ix, xi, xv, 31f, 46, 61, 68f, 77, 107, 110, 155, 167, 222f, 230, 234, 273  
 Lambalgen, M. van: 375, 378  
 Lambert, J. H.: 237, 241, 250, 273f, 277  
 Landau, E.: 235, 249, 251  
 Landin, P. J.: 356-358  
 Langley, P.: 400f  
 Laplace, P. S. M. de: 96, 198, 218, 254-256, 258, 260-263, 277, 309  
 Larmor, J.: 110  
 Laudan, L.: 49f, 59  
 Laugwitz, D.: 222  
 Lawvere, F. V.: xvi, 9, 30, 40, 46, 303, 312  
 Layzer, D.: 368, 378  
 Lebesgue, H.: 249  
 Leeuwen, J. van: 378  
 Lefschetz, S.: 83  
 Legendre, A. M.: 237-239, 242, 244-246, 248, 251-253, 258f, 277  
 Leibniz, G. W.: x, 15, 80, 86f, 96, 139, 194, 198, 200, 208, 210, 212, 217, 219, 222, 228, 297, 302f, 307, 309, 312  
 Lenin, V. I.: 15  
 Lerman, M.: 314, 347  
 Levin, L.: 373  
 Lèvy, A.: 332  
 Lindelöf, L.: 266f, 277  
 Lindgren, U.: 277  
 Lindsay, D.: viii  
 Littlewood, J. E.: 235, 249, 251  
 Lobachevski, N. I.: 310  
 Locke, J.: 230  
 Lommel, E.: 110  
 Lorentz, H. A.: 176, 179  
 Lorenzen, P.: 87  
 Louzil, J.: 311  
 Lovejoy, A. D.: 297, 307-309, 312  
 Loveland, D.: 373  
 Löwenheim, L.: 57  
 Lucas, E.: 231, 232, 233  
 Luce, R. D.: 77, 119, 121, 191  
 Ludwig, G.: 139, 146, 151, 156f, 164, 167  
 Luther, A.: 371, 378  
 Luxemburg, W. A. J.: 222, 225  
 Maass, W.: 338, 347  
 Mac Lane, S.: xiii, xvi, 12, 15, 17f, 23, 28, 30, 46, 87f, 126, 227, 350  
 Machover, M.: 332  
 Mackey, G. W.: 149  
 Maddy, P.: ix, xiii, xvi, 69, 77  
 Mahoney, M. S.: xi, xvi, 211f, 216f, 351  
 Mandelbrot, B.: 312  
 Manna, Z.: 361  
 Marcus, R. B.: 77, 401  
 Margenau, H.: 34-38, 44, 46  
 Markus, G.: 296, 312  
 Martin-Löf, P.: 70, 77, 187, 330, 332, 373  
 Maté, A.: 59  
 Matijašević, Yu.: 191, 320  
 Mauthner, F.: 15  
 Maxwell, J. C.: 15, 18, 136, 168, 175f, 376  
 May, R. M.: 97, 110  
 Mayberry: 17  
 Mayr, D.: 50, 59  
 McCarthy, J.: 353-358, 362  
 McCord, G. S.: 131  
 McCulloch, W. S.: 351, 353  
 McNaughton, R.: 351

- Mealy, G. H.: 351  
 Meissel, E.: 244, 247, 249, 251  
 Mendelsohn, E.: 210  
 Merriman, M.: 256, 258, 260, 264, 277f  
 Merz, J. T.: 301, 313  
 Meschkowski, H.: 224  
 Meyer, R.: 344, 361  
 Mill, J. S.: 59, 61, 63, 72-74, 77, 195, 308  
 Ming, L.: 373, 378  
 Minsky, M.: 350  
 Mises, L. von: 372  
 Mitchell, J.: 341f, 347  
 Mondesir, P. de: 244, 247  
 Montague, R.: 334  
 Moore, E. F.: 14, 351, 353  
 Mordell, L.: 11  
 Mormann, Th.: xv, 404  
 Moschovakis, Y.: 331, 333f, 336f, 342, 346f  
 Mosterin, J.: xv, 167  
 Mostowski, A.: 155  
 Moulines, C. U.: vii, xiv, 35-37, 45, 50, 58f, 76, 156, 167, 401, 403, 408, 410  
 Muchnik, A. A.: 325f, 333  
 Muldowney, P.: 191  
 Mundy, B.: 33, 42, 46  
 Mussolini, B.: 15  
 Nagel, E.: 31, 46, 49f, 52, 59  
 Naishtat, F. S.: 167  
 Nambu, Y.: 377  
 Narens, L.: 191  
 Naur, P.: 349  
 Neisser, U.: 38, 46  
 Neovius, V.: 266  
 Nerode, A.: 358  
 Nesbit, R. E.: 401  
 Neumann, J. von: 99, 144, 155, 168, 338  
 Newell, A.: 350  
 Newton, I.: 6, 94, 96, 135f, 138, 197, 208, 210-217, 219f, 228, 285, 295, 307, 362, 376  
 Nicolacopoulos, P.: 59  
 Nieuwentijdt, B.: 216f  
 Niiniluoto, I.: xv, 49f, 59, 61f, 64, 66, 69f, 72, 74f, 77, 302, 313  
 Noether, E.: 12  
 Nomizu, K.: 191  
 Nordheim, L.: 144  
 Normann, D.: 333  
 Novikoff, S. P.: 320  
 Nowak, L.: 50, 59  
 O'Hara, J.: 90  
 O'Hear, A.: 64, 77  
 Odifreddi, P.: 314, 317, 326f, 330, 338, 347  
 Ohm, G. S.: 98f, 110  
 Ohm, M.: 281, 289-291, 295  
 Olby, R. C.: 351  
 Oldenburg, H.: 214  
 Otte, M.: x, 87, 298, 303, 312f  
 Parmenides: 192  
 Parsons, Ch.: 65, 67, 77  
 Pascal, B.: 194  
 Peacock, G.: 103  
 Peano, G.: 303  
 Pearce, D.: 49-52, 59  
 Peirce, Ch. S.: 32, 39, 46, 95, 101, 260, 278, 308, 313  
 Peng-Yee, L.: 191  
 Penrose, R.: 368, 378  
 Perry, T. A.: 78  
 Peters, C. A. F.: 276  
 Petry, R. J.: 295  
 Pfaff, J. F.: 285, 295  
 Pindar: 132  
 Pitts, W.: 351, 353  
 Pizzetti, P.: 255f, 258, 260, 264, 269, 272-274, 278  
 Platek, R. A.: 330, 332, 334-336, 347  
 Plato: x, xii, 9, 59, 192-195, 297  
 Plotinus: 194  
 Plotkin, G.: 342, 347  
 Poincaré, H.: 11, 40f, 176  
 Poisson, S. D.: 96f, 198, 201, 256  
 Polanyi, M.: 81  
 Polos, L.: 59  
 Polya, G.: 61, 69, 77, 107f, 110, 231, 233f  
 Poncelet, J. V.: 80, 95f, 308  
 Popper, K. R.: 49, 59, 62, 64, 66, 68, 77, 91, 110, 393



- Pörn, I.: 77  
 Post, E. L.: 314f, 321-325, 327, 333, 344, 347  
 Postnikov, M. M.: 14  
 Potthoff, K.: 158, 165, 167  
 Prade, H.: 201  
 Putnam, H.: 32, 46, 59, 61f, 65, 70, 76f, 115f, 118-122, 196, 230, 320  
 Pycior, H.: 94, 110  
 Pythagoras: 17, 24, 154, 192, 206  
 Quine, W. V. O.: 9, 31f, 44, 46, 56, 59, 115f, 118-122, 197, 230  
 Rabin, M. O.: 352f  
 Radnitzky, G.: 46  
 Radon, J.: 7  
 Ramanujan, S.: 249  
 Ramsey, F. P.: 164, 199f  
 Randell, B.: 349  
 Rantala, V.: xv, 50-53, 57, 59  
 Raphson, J.: 215  
 Ravetz, J. R.: 109f  
 Reade, C.: 341, 347  
 Reed, K.: 131  
 Reichenbach, H.: 59, 77  
 Resnik, D. B.: XIII, xv, 123, 131  
 Resnik, M. D.: 32, 44, 46, 65, 77, 128  
 Reuschle, K. G.: 255f, 269f, 275, 278  
 Reynolds, J. C.: 361  
 Reynolds, R. G.: 397f, 401  
 Richardson, D.: 191  
 Riemann, G. F. B.: 7, 11, 219, 234, 243-249, 251  
 Ringen, J.: 74, 78  
 Ritchie, R. W.: 361  
 Ritt, J. F.: 12  
 Roberval, G. P.: 211, 213  
 Robinson, A.: 210, 221-228  
 Robinson, J.: 191, 320  
 Rochester, N.: 356  
 Rogers, H.: 325-327, 347  
 Rolle, M.: 208, 216f  
 Rootselaar, B. von: 311  
 Rorty, R.: 313  
 Rosser, J. B.: 354  
 Rostand, F.: 106f, 110  
 Rota, G.-C.: 296, 303f, 313  
 Rothe, H. A.: 285, 295  
 Rowe, D.: 110  
 Ruelle, D.: 88  
 Russell, B.: 9, 59, 79, 195  
 Sacks, G. E.: 327, 330-333, 346f  
 Sadovsky, V.: 59  
 Saint Augustine: 194  
 Salanskis, J. M.: 86  
 Schaffner, K. F.: 401  
 Schanuel, S.: 17, 25, 30  
 Scheibe, E.: xv, 144, 156-159, 161, 166f  
 Schelling, F. W. J.: 313  
 Schering, E.: 250  
 Scherk, J.: 377  
 Schiaparelli, G. V.: 271f, 278  
 Schilpp, P. A.: 46  
 Schlick, M.: 313  
 Schmidt, H.-J.: xv, 59, 164, 167, 227  
 Schneider, I.: 258, 278  
 Schnirelman, L.: 250  
 Scholz, H.: 221  
 Schooten, F.: 110  
 Schröder, E.: 95  
 Schumacher, H. C.: 258  
 Schur, I.: 7  
 Schwarz, J.: 377f  
 Scott, D. S.: 353, 358-360  
 Scriba, J.: 232  
 Seeliger, H. von: 274, 278  
 Seidel, Ph. L. von: 274, 278  
 Seifert, H.: 83  
 Selberg, A.: 236, 248, 250  
 Seligman, G.: 226  
 Shafer, G.: 201  
 Shakespeare, W.: 132  
 Shanks, D.: 234  
 Shannon, C.: 351, 353f, 373, 378  
 Shapiro, E. Y.: 401  
 Shapiro, S.: 122  
 Shenitzer, A.: 280  
 Shepherdson, J.: 315, 334f, 347  
 Sherp, D. H.: 313  
 Sheynin, O. B.: 259, 275, 278  
 Shoenfield, J. R.: 156, 164f, 167, 191

- Shore, R.: 331f, 347  
 Sibelius, J.: 65  
 Sieg, W.: 314, 339  
 Simon, D.: viii  
 Simon, H. A.: 313, 401  
 Simon, P.: 277  
 Simons, P. M.: 61, 78  
 Simpson, S.: 326f, 333, 347  
 Sinaceur, H.: 86  
 Skolem, Th.: 9  
 Slaman, T.: 338, 347  
 Sluse, R.: 213  
 Smith, B.: 78  
 Smith, J.: 244, 378  
 Smolin, L.: 191  
 Smullyan, R.: 336, 347  
 Sneed, J. D.: xiv, 35-37, 45, 50, 58f, 75f, 147, 156, 167, 294f, 401, 403, 408, 410  
 Soare, R.: 327, 347  
 Sokolowski, R.: 313  
 Solomonoff, R.: 372-374, 378  
 Solovay, R.: 344f  
 Spalt, D.: 81  
 Spanier, E. H.: 89  
 Spehr, F. W.: 282f, 285-291, 293, 295  
 Spinoza, B.: x, 194f  
 Spohn, W.: 201  
 Stasheff, J.: 7  
 Stearns, R. E.: 351f  
 Steenrod, N.: 83  
 Steffensen: 235  
 Stegmüller, W.: 59, 75, 78, 173, 191, 294f  
 Steinbring, H.: 313  
 Steiner, J.: 17, 24  
 Steinheil, C. A. von: 274, 278  
 Steinitz: 12  
 Stenius, E.: 59, 67f, 78  
 Sterling, L.: 401  
 Sternberg, S.: 191  
 Stevens, S. S.: 33  
 Stigler, St. M.: 254, 258, 278  
 Stockmeyer: 344  
 Stokes, G. G.: 244  
 Stone, E. J.: 271f, 278  
 Strachey, C.: 349, 354, 357f, 360f  
 Strawson, P. F.: 313  
 Sturgis, H.: 315, 334, 347  
 Sullivan, K.: 227  
 Suphan, B.: 294  
 Suppe, F.: 31, 46, 49, 55, 59  
 Suppes, P.: 33, 46, 59, 71, 75, 77f, 119, 121, 147, 156, 167, 169f, 172, 174, 190f  
 Tait, P. G.: 258, 262, 278  
 Tait, W. W.: 65, 78  
 Takeuti, G.: 153, 332  
 Tamburrini, G.: 317, 348  
 Tannéry, P.: 85  
 Tarski, A.: 59, 61, 150  
 Taubes, C. H.: 191  
 Taylor, B.: 218, 286-288  
 Thagard, P. R.: 401  
 Theaetetus: 193  
 Thibaut, B. F.: 283  
 Thiébaud, N.: 25, 30  
 Thomas, W.: 156, 159, 161, 167  
 Thomson, W.: 244  
 Threlfall, W.: 83  
 Tierney, M.: 12  
 Tilly, J. M. de: 262, 267, 271-273, 279  
 Titshmarsh, E. C.: 250  
 Toraldo di Francia, G.: 191  
 Torelli, G.: 235  
 Torretti, R.: xv  
 Troelstra, A. S.: 63, 78, 331, 348  
 Truesdell, C.: 205  
 Tucker, J.: 342, 348  
 Tuomela, R.: 50, 59, 77f, 301, 313  
 Turing, A.: 314-321, 324f, 327f, 330, 332-335, 338-340, 343-345, 348, 351, 353, 361, 371, 378  
 Turnbull, H. W.: 295  
 Tversky, A.: 77, 119, 121, 191  
 Tymoczko, T.: ix, xvi, 61, 66, 74, 77f, 107, 110, 231  
 Ullman, J. D.: 338, 345, 361, 402  
 Urbach, P. M.: 200f  
 Vallée Poussin, Ch. de la: 236, 248f  
 Van Fraassen, B. C.: 36, 46  
 Van t'Hooft, V.: 377  
 Veblen, O.: 83

- Vega, G.: 238, 241f, 249-251  
Veneziano, G.: 377  
Vercelloni, L.: 104, 110  
Viète, F.: 85, 94, 110, 211  
Vitanyi, P.: 373, 378  
Vivanti, G.: 224  
Waerden, B. L. van der: 12, 194, 201, 279  
Wahsner, R.: 300, 311  
Weaver, W.: 378  
Weber, B.: 226f, 378  
Weber, E. H.: 274  
Weierstrass, K.: 219f, 223, 285, 292  
Weil, A.: 231f  
Weinberg, S.: 116-118, 120, 122f, 125-127, 141-144  
Weingartner, P.: 77, 401  
Weischedel, W.: 295  
Weiss, P.: 46  
West, B. J.: 100, 111  
Wexelblat, R.: 356  
Weyl, H.: 7, 155, 300, 313  
Whitehead, A. N.: 9, 351  
Whiteside, D. T.: 214  
Wiener, N.: 367, 378  
Wightman, A. S.: 191  
Wigner, E.: 105, 111, 141, 300  
Williams, H. C.: 231  
Winograd, S.: 361  
Winter, E.: 311  
Wise, M. N.: 99, 111  
Wittgenstein, L.: 82, 195  
Wolff, M.: 293, 295  
Worrall, J.: 167, 222  
Wussing, H.: 41, 46  
Young, Th.: 264, 279  
Youschkevitch, A. P.: 310, 313  
Yuxin, Z.: 31, 46  
Zellweger, S.: 95, 111  
Zermelo, E.: 9, 147, 156, 159, 168, 173, 196  
Zucker, J.: 342, 348  
Zytkow, J. M.: 401



**Walter de Gruyter**  
**Berlin · New York**

**GRUNDLAGEN DER KOMMUNIKATION  
UND KOGNITION  
FOUNDATIONS OF COMMUNICATION  
AND COGNITION**

---

**Meaning Scepticism**

*Edited by Klaus Puhl*

Groß-Oktav. IX, 258 Seiten. 1991. Ganzleinen DM 142,-  
ISBN 3 11 011833 5

**RALF TWENHÖFEL**

**Wissenschaftliches Handeln**

**Aspekte und Bestimmungsgründe der Forschung**

Groß-Oktav. XI, 303 Seiten. 1991. Ganzleinen DM 148,-  
ISBN 3 11 012416

**RALPH AXEL MÜLLER**

**Der (un)teilbare Geist**

**Modularismus und Holismus in der  
Kognitionsforschung**

Oktav. XIX, 443 Seiten. 1991. Gebunden DM 198,-  
ISBN 3 11 012916 7

Preisänderungen vorbehalten



**Walter de Gruyter**  
**Berlin · New York**

**GRUNDLAGEN DER KOMMUNIKATION  
UND KOGNITION  
FOUNDATIONS OF COMMUNICATION  
AND COGNITION**

---

LORENZ BRUNO PUNTEL

**Grundlagen einer Theorie der  
Wahrheit**

Groß-Oktav. XIII, 408 Seiten. 1990. Ganzleinen DM 198,-  
ISBN 3 11 012079 8

HELMUT SCHNELLE

**Die Natur der Sprache**

**Die Dynamik der Prozesse des Sprechens und  
Verstehens**

Oktav. XII, 671 Seiten. 1991. Gebunden DM 258,-  
ISBN 3 11 012704 0

**Emergence or Reduction?**

**Essays on the Prospects of Nonreductive Physicalism**

*Edited by Ansgar Beckermann, Hans Flohr, Jaegwon Kim*

1992. Octavo. VIII, 315 pages, 8 illustrations.  
Cloth DM 152,- ISBN 3 11 012880 2

Preisänderungen vorbehalten